# Asymptotic MSE Performance of ML Direction-Of-Arrival Estimation with Estimated Noise Covariance

**Christ D. Richmond***

**Foundations of Computational Mathematics**

**Workshop on Random Matrix Theory, Algorithms, and Applications**
**July 12-14th 2011, Budapest, Hungary**



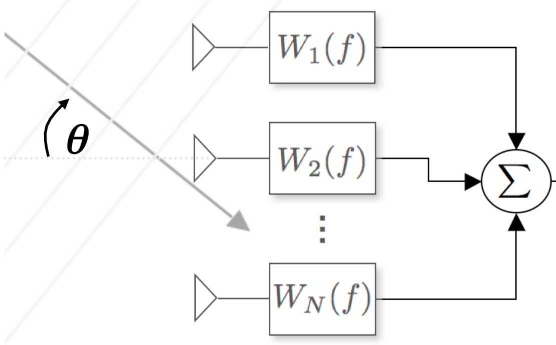
*The author is a Senior Staff at MIT Lincoln Laboratory, Lexington, MA and Senior Member of the IEEE.

# Outline

- Introduction
  - Adaptive sensor arrays
  - Maximum-Likelihood
- Mean-squared error prediction
- Numerical examples
- Summary

# Adaptive Sensor Arrays

## General Beamformer



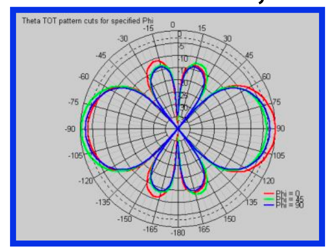
$$y(t) = \sum_n W_n \cdot x_n(t) = \mathbf{w}^H \mathbf{x}$$

$$\mathrm{cov}(\mathbf{x}) = \mathbf{R} \quad E\{\mathbf{x}\} = S\mathbf{v}$$

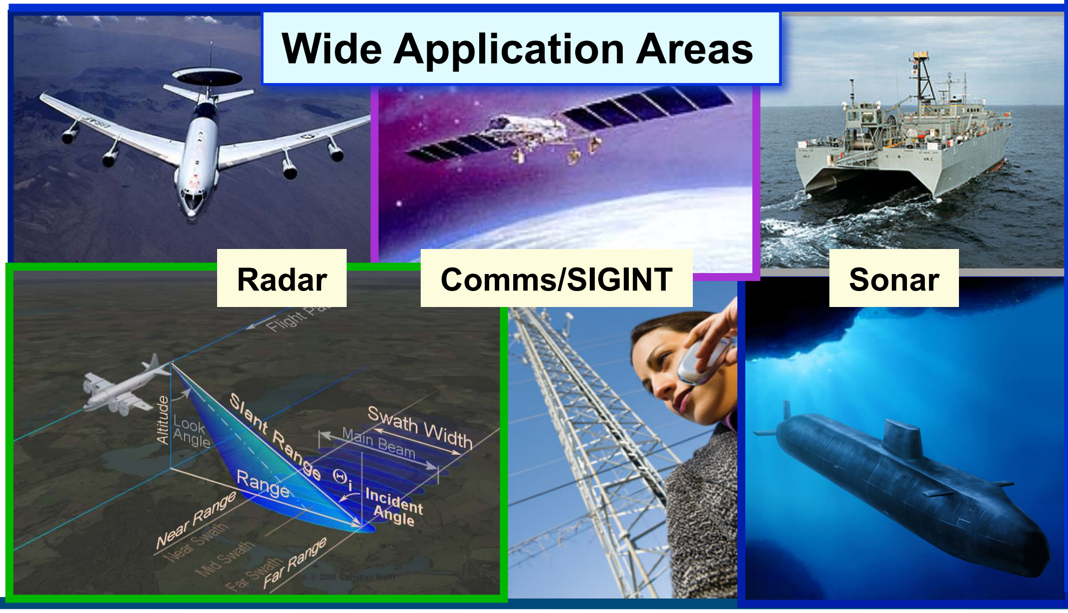
- Spatial filtering
- Adapt to changing environment
- Spatial diversity increases capacity

## Conventional System Design

- Array topology / # sensors
  - Resolution, sidelobes, ambiguities, etc.
- Operating frequency / bandwidth
- Power constraints, cost, etc.



## Wide Application Areas




## Algorithm Design/Analysis

- Detection, estimation / localization / classification, and tracking
- Metrics: Optimal Neyman-Pearson, Max signal-to-interference + noise ratio (SINR), Min mean squared error (MSE), etc.

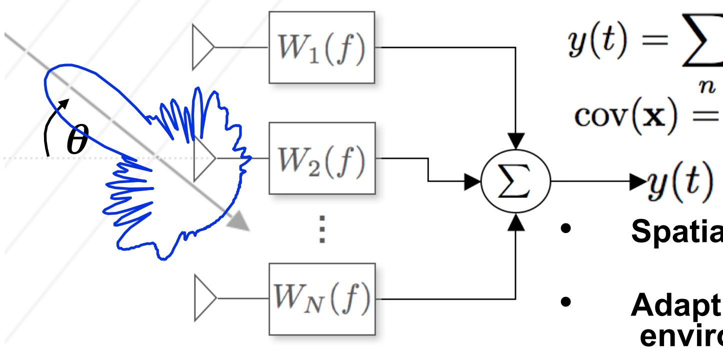
$$\boxed{\mathbf{w} \propto \mathbf{R}^{-1}\mathbf{v}}$$

*Weiner Solution*

- Theoretical performance analysis

# Adaptive Sensor Arrays

## General Beamformer



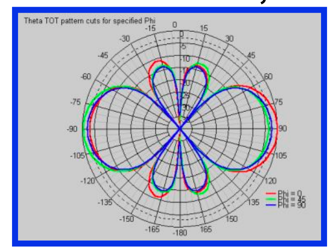

$$y(t) = \sum_n W_n \cdot x_n(t) = \mathbf{w}^H \mathbf{x}$$

$$\mathrm{cov}(\mathbf{x}) = \mathbf{R} \quad E\{\mathbf{x}\} = S\mathbf{v}$$

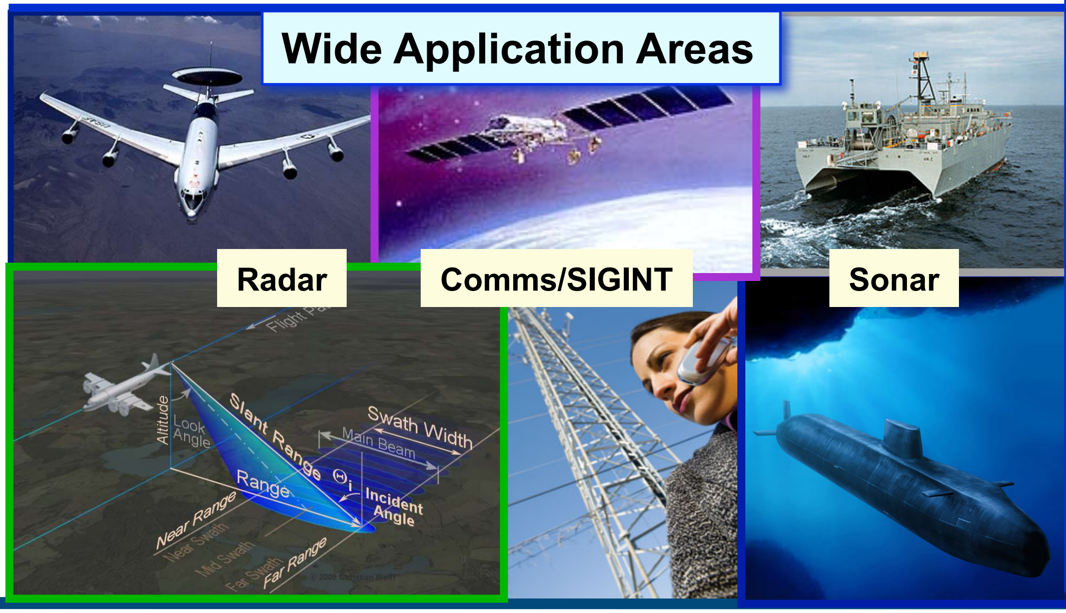
- Spatial filtering
- Adapt to changing environment
- Spatial diversity increases capacity

## Conventional System Design

- Array topology / # sensors
  - Resolution, sidelobes, ambiguities, etc.
- Operating frequency / bandwidth
- Power constraints, cost, etc.




## Wide Application Areas




## Algorithm Design/Analysis

- Detection, estimation / localization / classification, and tracking
- Metrics: Optimal Neyman-Pearson, Max signal-to-interference + noise ratio (SINR), Min mean squared error (MSE), etc.

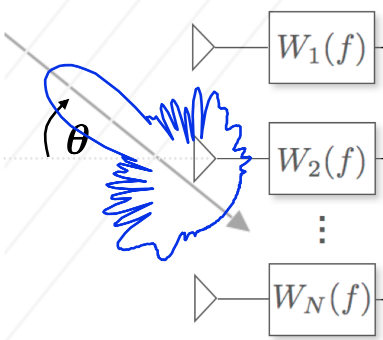
$$\boxed{\mathbf{w} \propto \mathbf{R}^{-1}\mathbf{v}}$$

*Weiner Solution*

- Theoretical performance analysis

# Adaptive Sensor Arrays

## General Beamformer



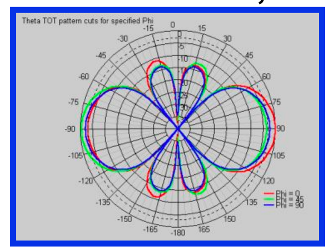

$$y(t) = \sum_n W_n \cdot x_n(t) = \mathbf{w}^H \mathbf{x}$$

$$\mathrm{cov}(\mathbf{x}) = \mathbf{R} \quad E\{\mathbf{x}\} = S\mathbf{v}$$

- Spatial filtering
- Adapt to changing environment
- Spatial diversity increases capacity

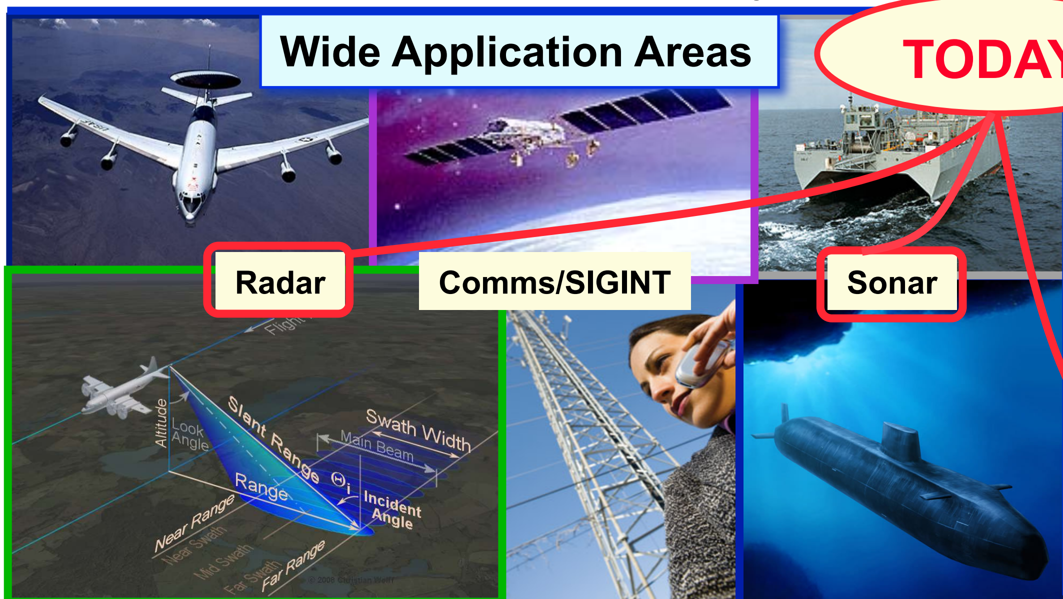
## Conventional System Design

- Array topology / # sensors
  - Resolution, sidelobes, ambiguities, etc.
- Operating frequency / bandwidth
- Power constraints, cost, etc.




## Wide Application Areas

TODAY




## Algorithm Design/Analysis

- Detection, estimation / localization / classification, and tracking
- Metrics: Optimal Neyman-Pearson, Max signal-to-interference + noise ratio (SINR), Min mean squared error (MSE), etc.

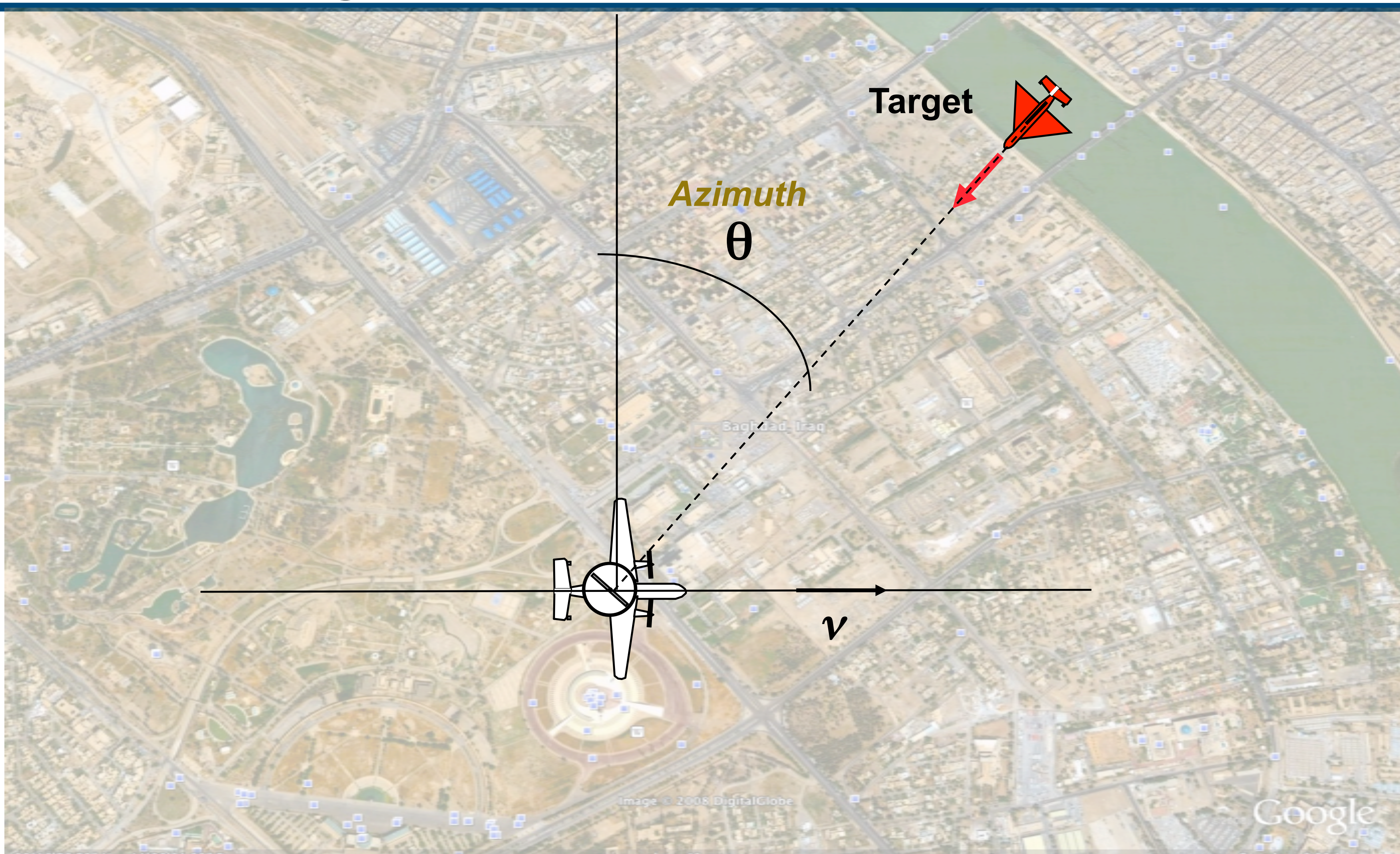
$$\mathbf{w} \propto \mathbf{R}^{-1}\mathbf{v}$$

*Weiner Solution*

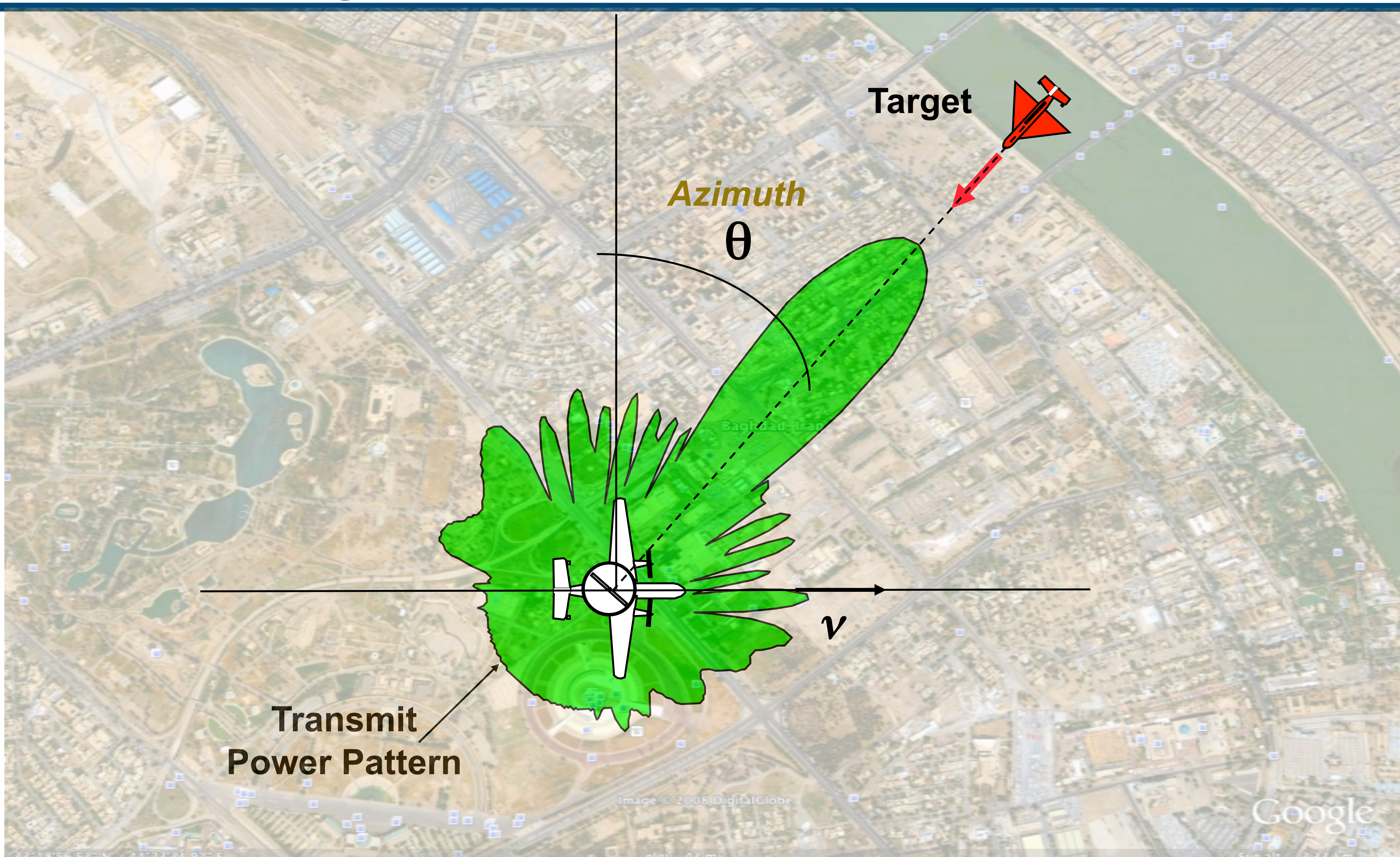
- Theoretical performance analysis

# Airborne Surveillance Radars: Signal and Interference Environment

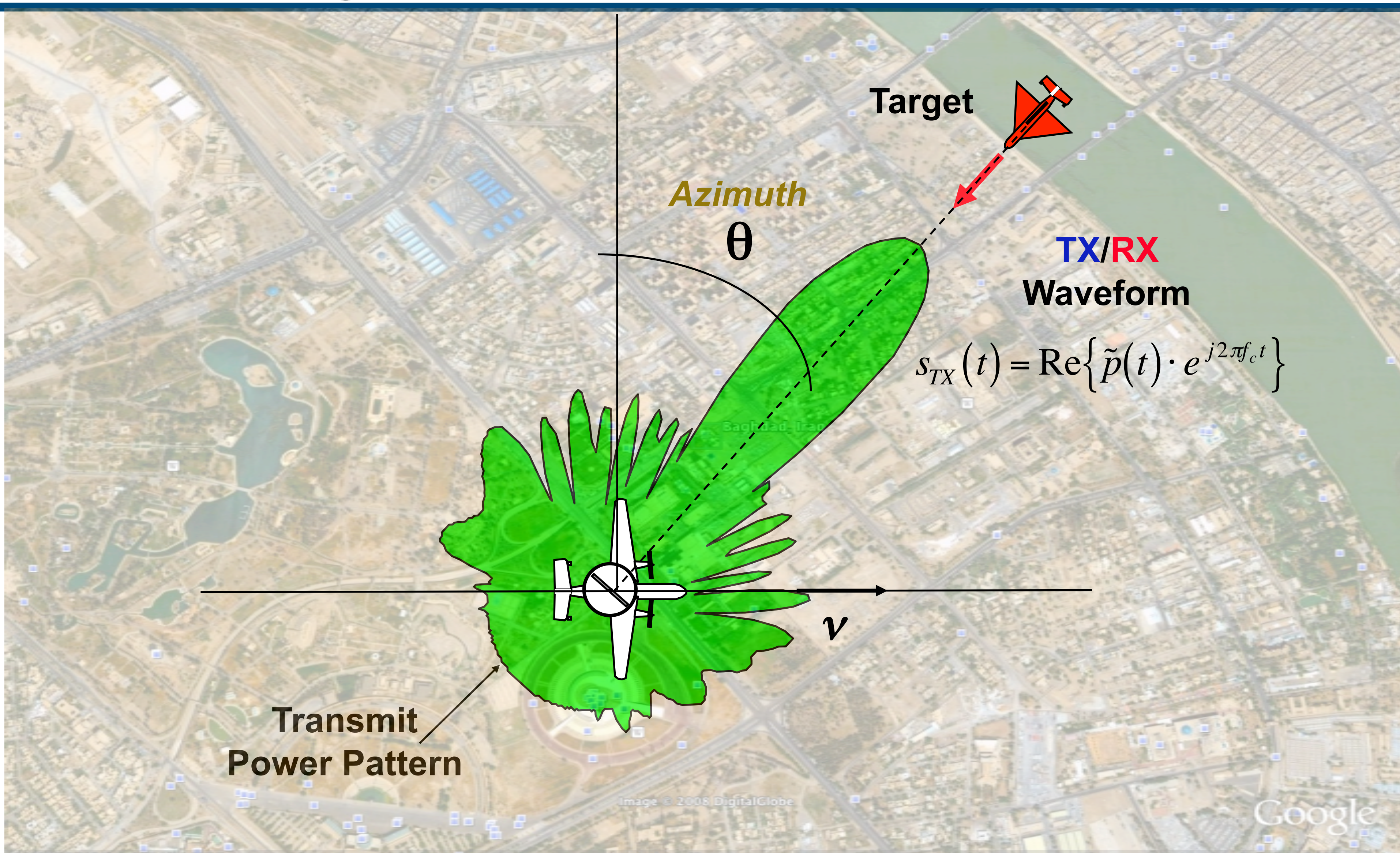
# Airborne Surveillance Radars: Signal and Interference Environment



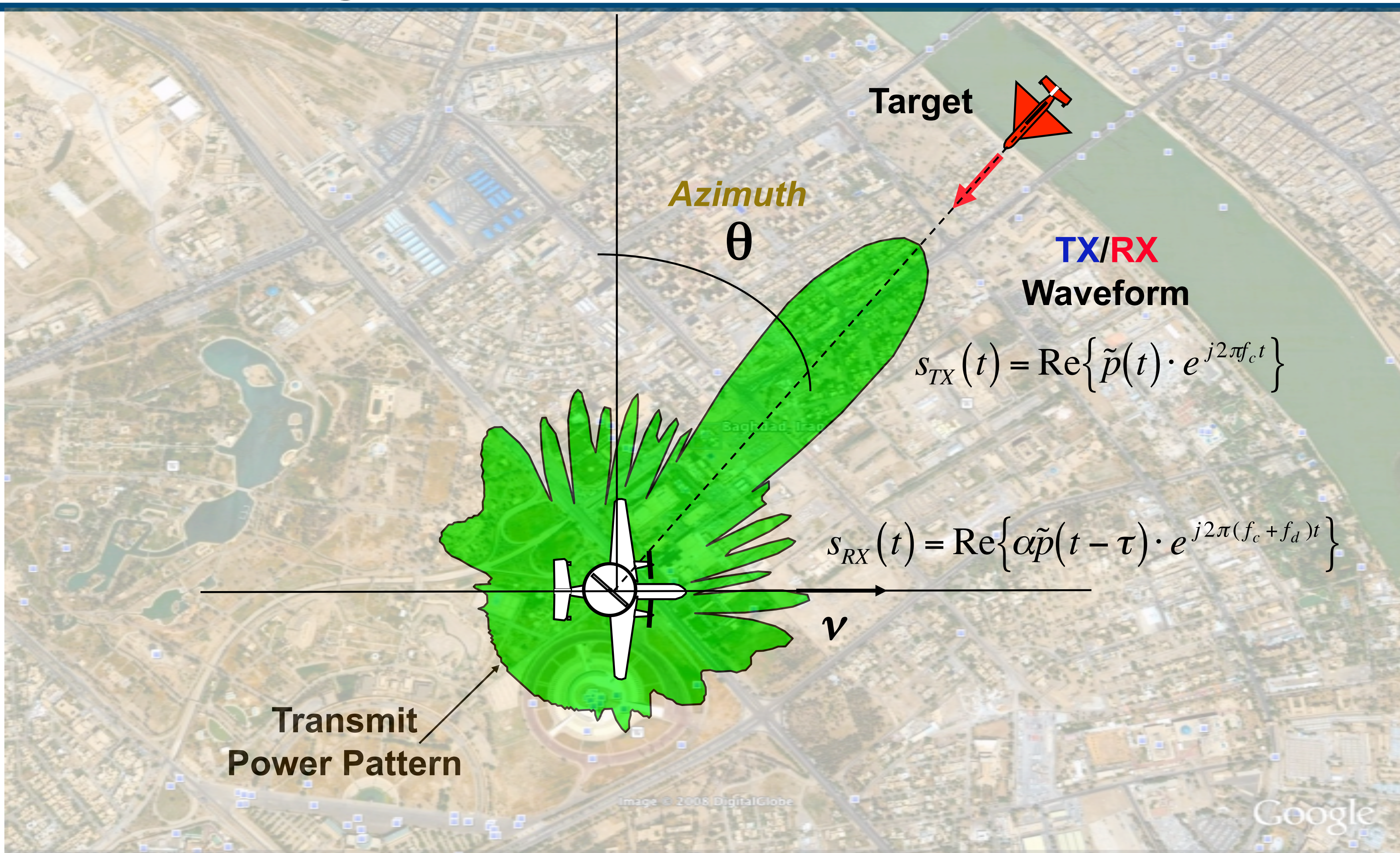
# Airborne Surveillance Radars: Signal and Interference Environment



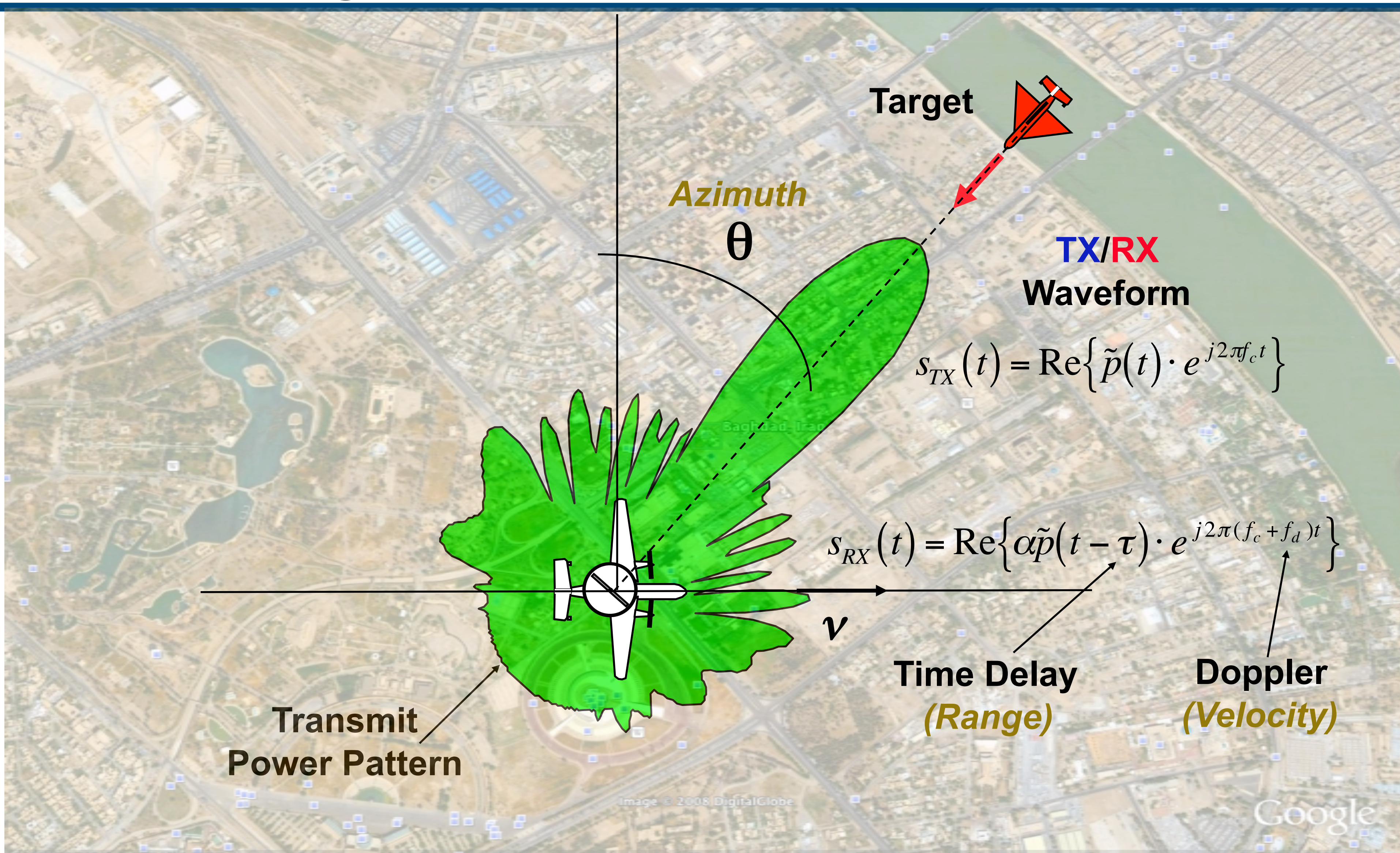
# Airborne Surveillance Radars: Signal and Interference Environment



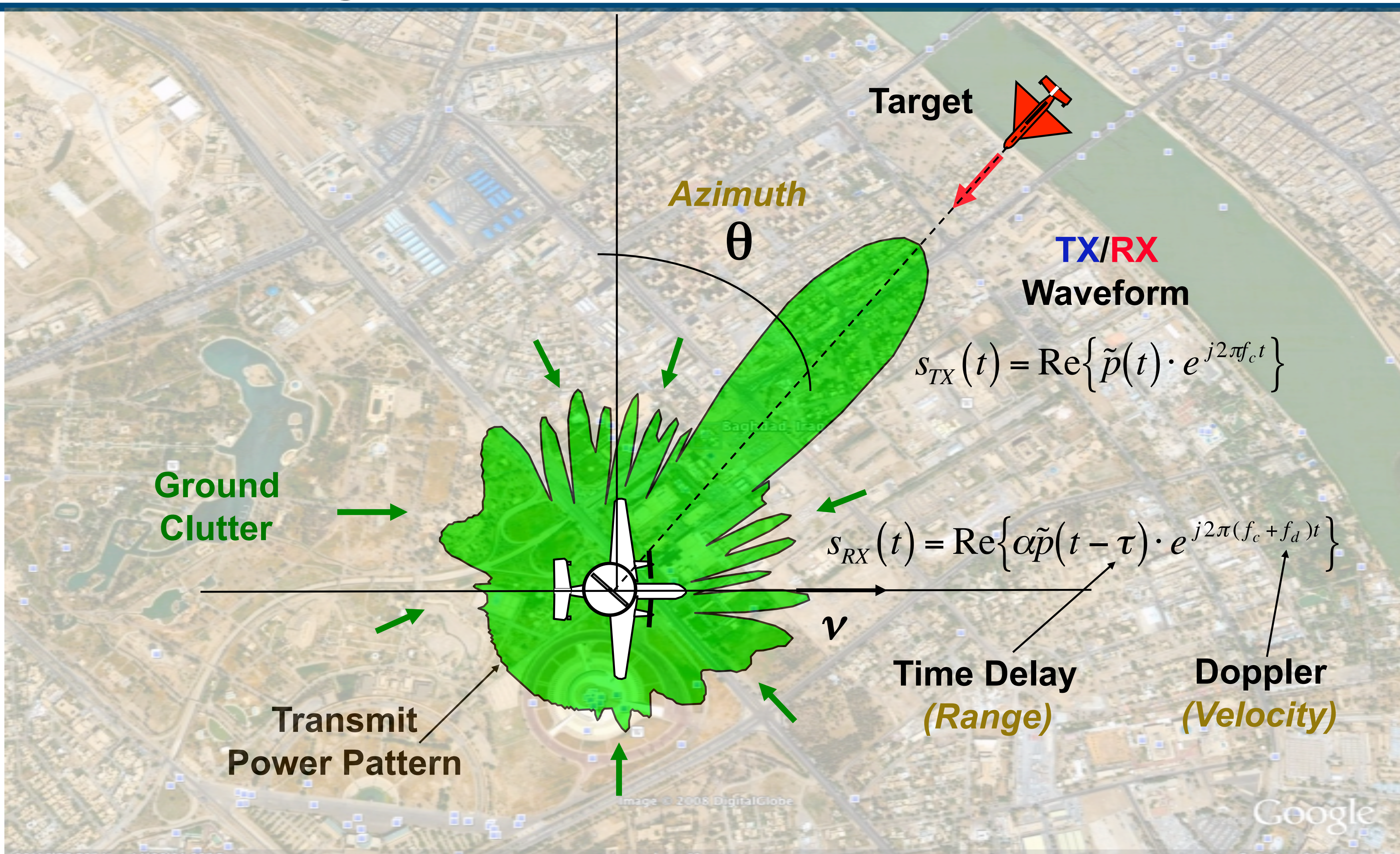
# Airborne Surveillance Radars: Signal and Interference Environment



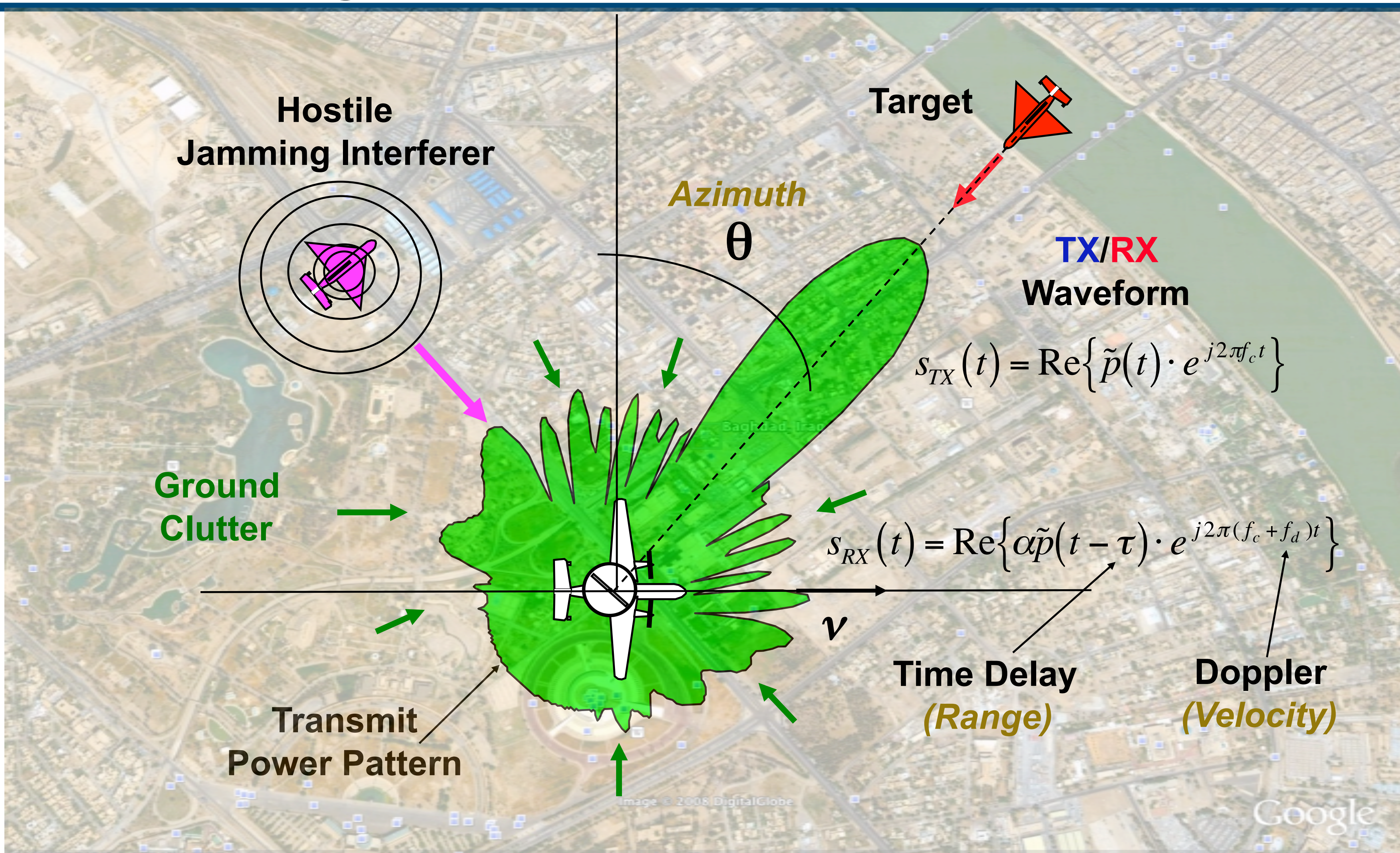
# Airborne Surveillance Radars: Signal and Interference Environment



# Airborne Surveillance Radars: Signal and Interference Environment



# Airborne Surveillance Radars: Signal and Interference Environment



# Maximum-Likelihood Estimation

- **Post-detection it is desired to estimate parameters of target**

- **Maximum-likelihood (ML) often effective approach:**

$$\widehat{\theta}_{ML} = \arg\max_{\theta} \ln p(\mathbf{x}|\theta, \mathbf{R}) \qquad \widehat{\theta}_{ML} = \arg\max_{\theta, \mathbf{R}} \ln p(\mathbf{x}, \mathbf{x}(1), \ldots, \mathbf{x}(L)|\theta, \mathbf{R})$$

# Maximum-Likelihood Estimation

- **Post-detection it is desired to estimate parameters of target**

- **Maximum-likelihood (ML) often effective approach:**

Test cell

Noise only samples

$$\widehat{\theta}_{ML} = \arg\max_{\theta} \ln p(\mathbf{x}|\theta, \mathbf{R}) \qquad \widehat{\theta}_{ML} = \arg\max_{\theta,\mathbf{R}} \ln p(\mathbf{x}, \mathbf{x}(1), \ldots, \mathbf{x}(L)|\theta, \mathbf{R})$$

# Maximum-Likelihood Estimation

- **Post-detection it is desired to estimate parameters of target**

- **Maximum-likelihood (ML) often effective approach:**

Test cell

Noise only samples

$$\widehat{\theta}_{ML} = \arg\max_{\theta} \ln p(\mathbf{x}|\theta, \mathbf{R}) \qquad \widehat{\theta}_{ML} = \arg\max_{\theta, \mathbf{R}} \ln p(\mathbf{x}, \mathbf{x}(1), \ldots, \mathbf{x}(L)|\theta, \mathbf{R})$$

| Clairvoyant Matched Filter | **R known:** | **R unknown:** | Adaptive Matched Filter |
|---|---|---|---|
| | $\widehat{\theta}_{ML} = \arg\max_{\theta} \dfrac{\lvert\mathbf{v}^H(\theta)\mathbf{R}^{-1}\mathbf{x}\rvert^2}{\mathbf{v}^H(\theta)\mathbf{R}^{-1}\mathbf{v}(\theta)}$ | $\widehat{\theta}_{ML} = \arg\max_{\theta} \dfrac{\lvert\mathbf{v}^H(\theta)\widehat{\mathbf{R}}^{-1}\mathbf{x}\rvert^2}{\mathbf{v}^H(\theta)\widehat{\mathbf{R}}^{-1}\mathbf{v}(\theta)}$ | |

**where** $\mathbf{x} \sim \mathcal{CN}_N\left(S\mathbf{v}(\theta_1), \mathbf{R}\right)$, $\widehat{\mathbf{R}} \propto \sum_{l=1}^{L} \mathbf{x}(l)\mathbf{x}^H(l)$, **and** $\mathbf{x}(l) \sim \mathcal{CN}_N\left(\mathbf{0}, \mathbf{R}\right)$

# Maximum-Likelihood Estimation

- **Post-detection it is desired to estimate parameters of target**

- **Maximum-likelihood (ML) often effective approach:**

**Test cell**

**Noise only samples**

$$\widehat{\theta}_{ML} = \arg\max_{\theta} \ln p(\mathbf{x}|\theta, \mathbf{R}) \qquad \widehat{\theta}_{ML} = \arg\max_{\theta,\mathbf{R}} \ln p(\mathbf{x}, \mathbf{x}(1), \ldots, \mathbf{x}(L)|\theta, \mathbf{R})$$

| ***Clairvoyant Matched Filter*** | **<u>R known:</u>** | **<u>R unknown:</u>** | ***Adaptive Matched Filter*** |
|---|---|---|---|
| | $\widehat{\theta}_{ML} = \arg\max_{\theta} \dfrac{\lvert\mathbf{v}^H(\theta)\mathbf{R}^{-1}\mathbf{x}\rvert^2}{\mathbf{v}^H(\theta)\mathbf{R}^{-1}\mathbf{v}(\theta)}$ | $\widehat{\theta}_{ML} = \arg\max_{\theta} \dfrac{\lvert\mathbf{v}^H(\theta)\widehat{\mathbf{R}}^{-1}\mathbf{x}\rvert^2}{\mathbf{v}^H(\theta)\widehat{\mathbf{R}}^{-1}\mathbf{v}(\theta)}$ | |

**where** $\mathbf{x} \sim \mathcal{CN}_N(S\mathbf{v}(\theta_1), \mathbf{R})$, $\widehat{\mathbf{R}} \propto \sum_{l=1}^{L} \mathbf{x}(l)\mathbf{x}^H(l)$, **and** $\mathbf{x}(l) \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{R})$

- **ML requires non-linear search ("beamsplitting")**
  - **Optimal asymptotically, but with threshold effect**
- ***GOAL: Predict MSE performance near threshold***

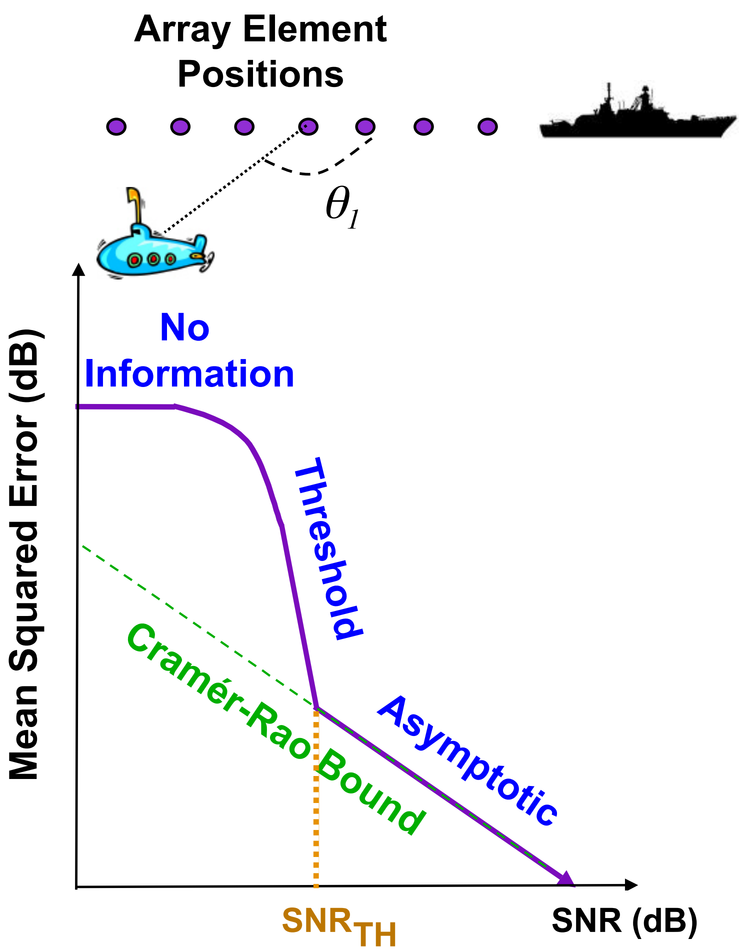
***Mean Squared Error (MSE):***

$$E\left\{\left(\widehat{\theta}_{ML} - \theta_1\right)^2\right\}$$

# Mean-Squared Error Performance: No Mismatch vs Mismatch

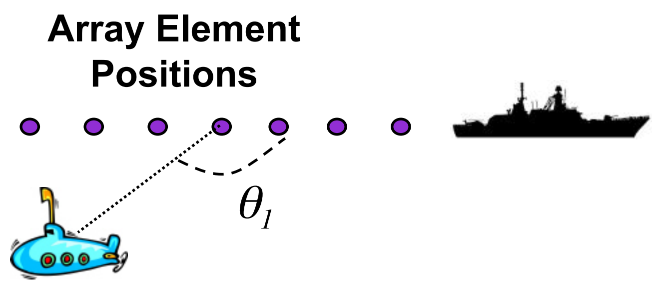
$$\hat{\theta}_{ML} = \arg\max_{\theta} \; t_{ML}(\theta,\text{data})$$

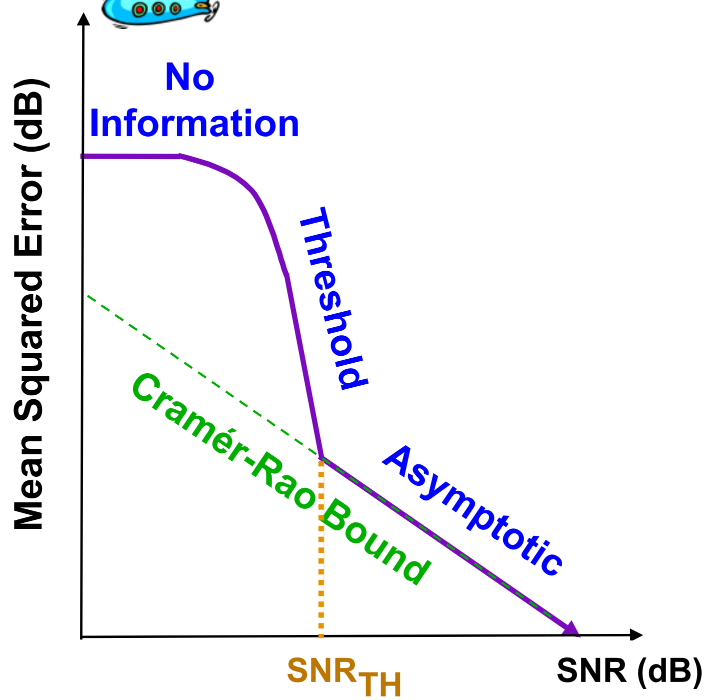
## No Mismatch



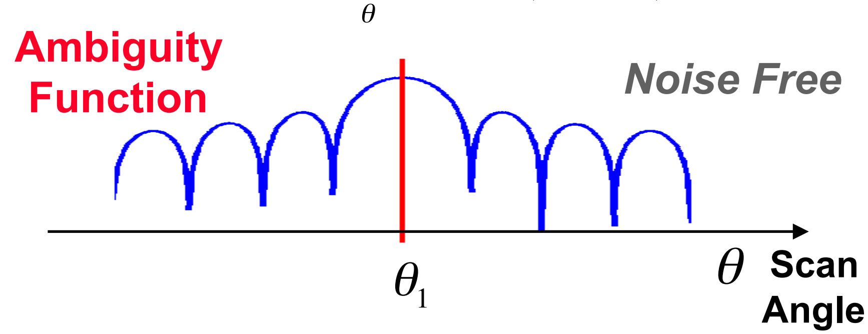
# Mean-Squared Error Performance: No Mismatch vs Mismatch

## No Mismatch



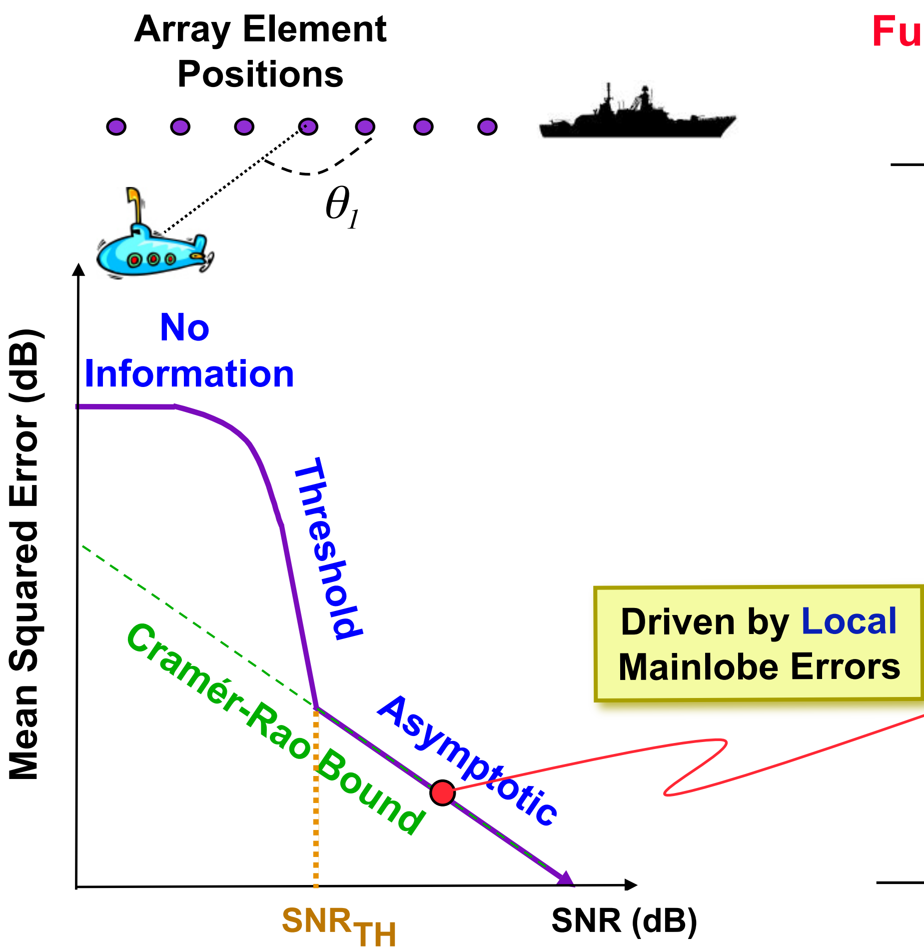

$$\hat{\theta}_{ML} = \arg\max_{\theta} \; t_{ML}(\theta,\text{data})$$

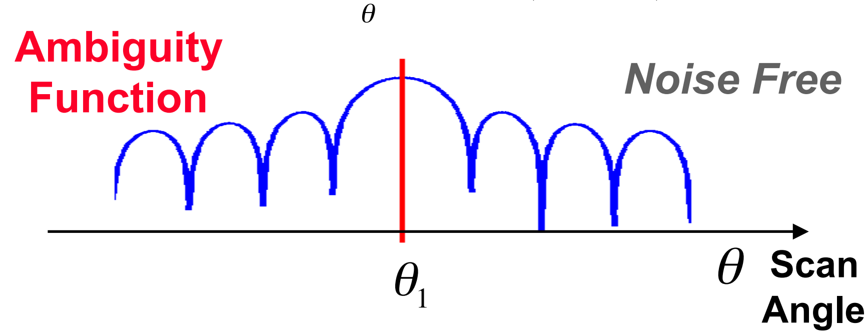


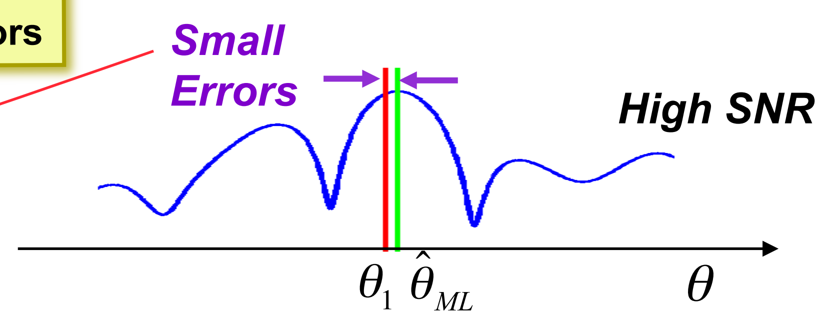
# Mean-Squared Error Performance: No Mismatch vs Mismatch

## No Mismatch

$$\hat{\theta}_{ML} = \arg\max_{\theta}\ t_{ML}(\theta,\text{data})$$

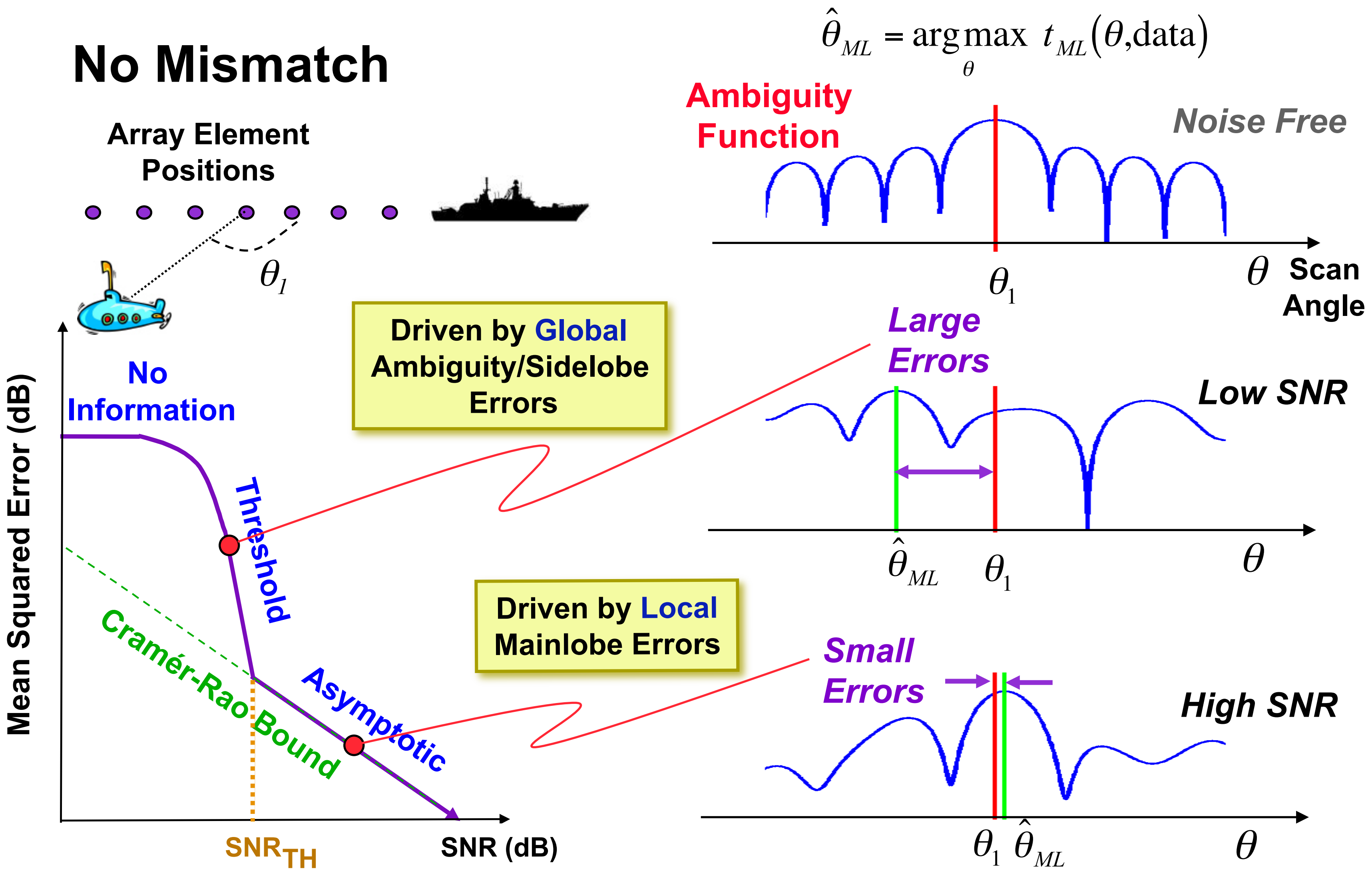


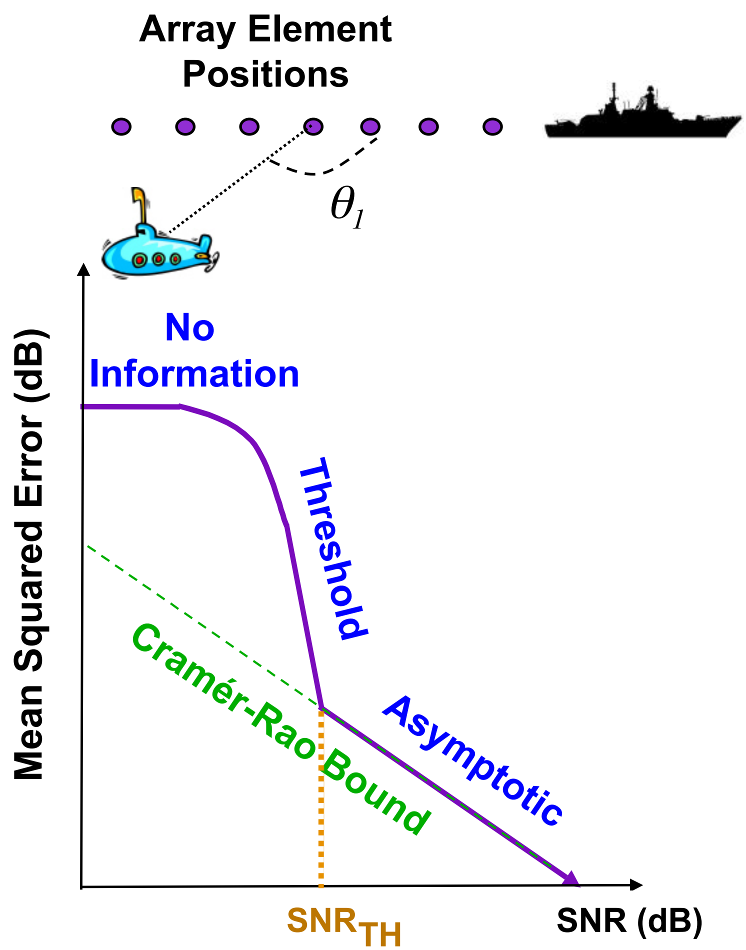
# Mean-Squared Error Performance: No Mismatch vs Mismatch

## No Mismatch

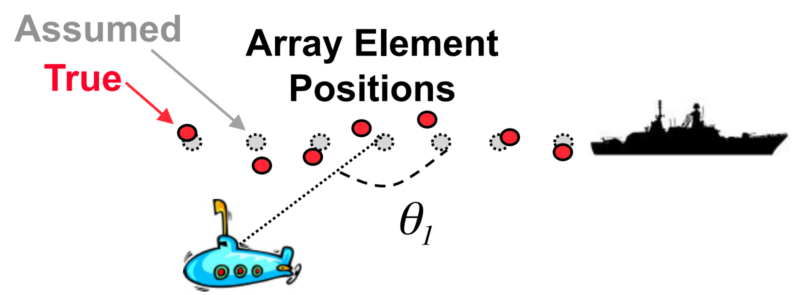


# Mean-Squared Error Performance: No Mismatch vs Mismatch

## No Mismatch

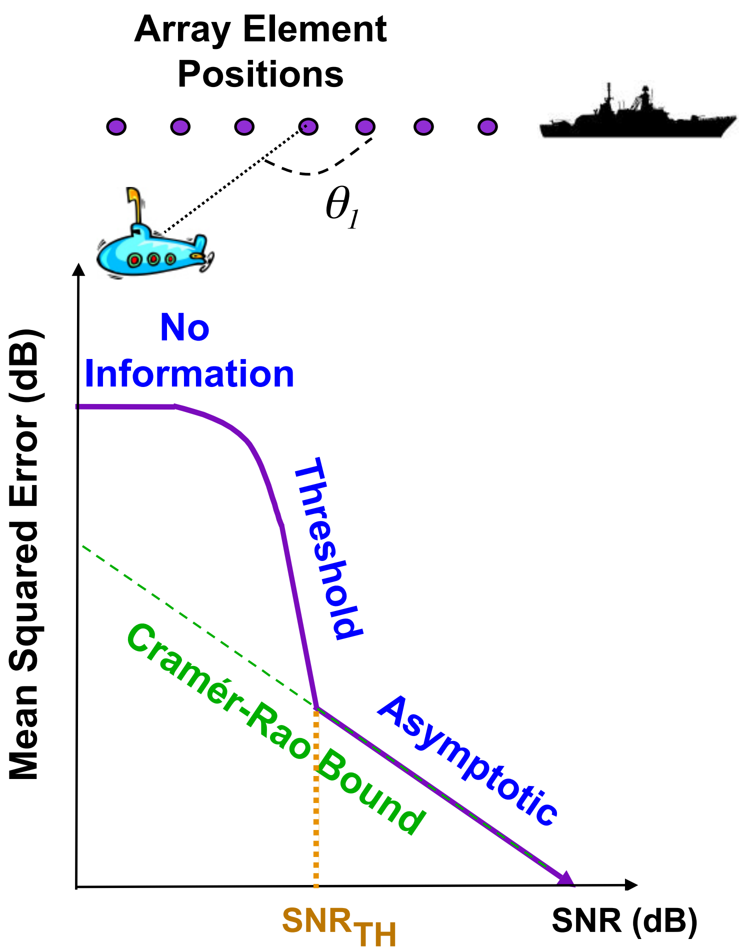



## Signal Mismatch

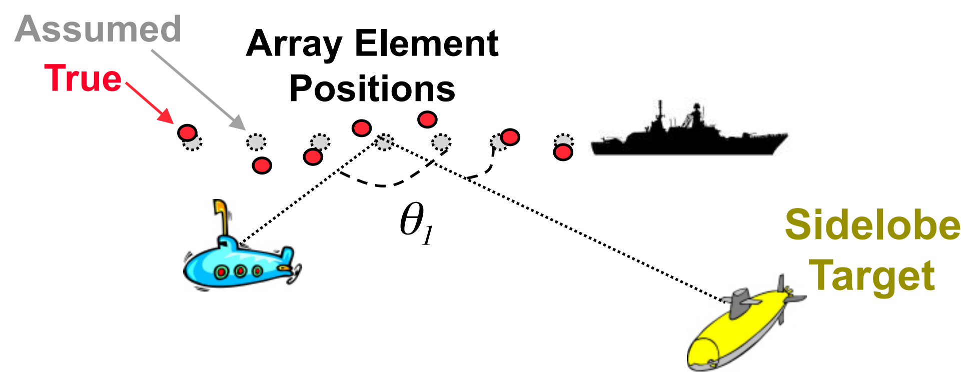


# Mean-Squared Error Performance: No Mismatch vs Mismatch

## No Mismatch

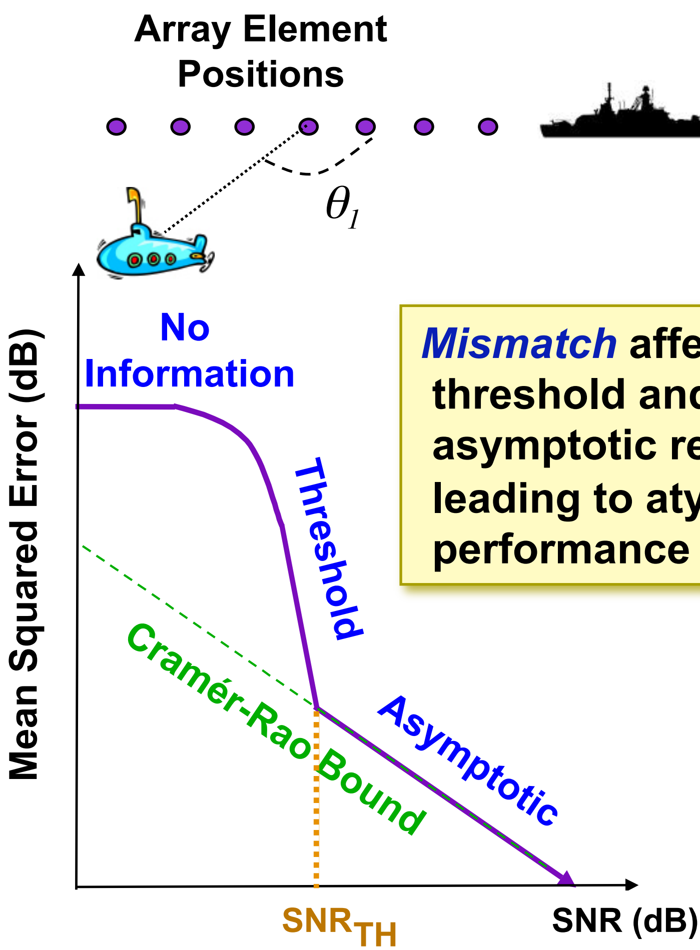



## Signal Mismatch



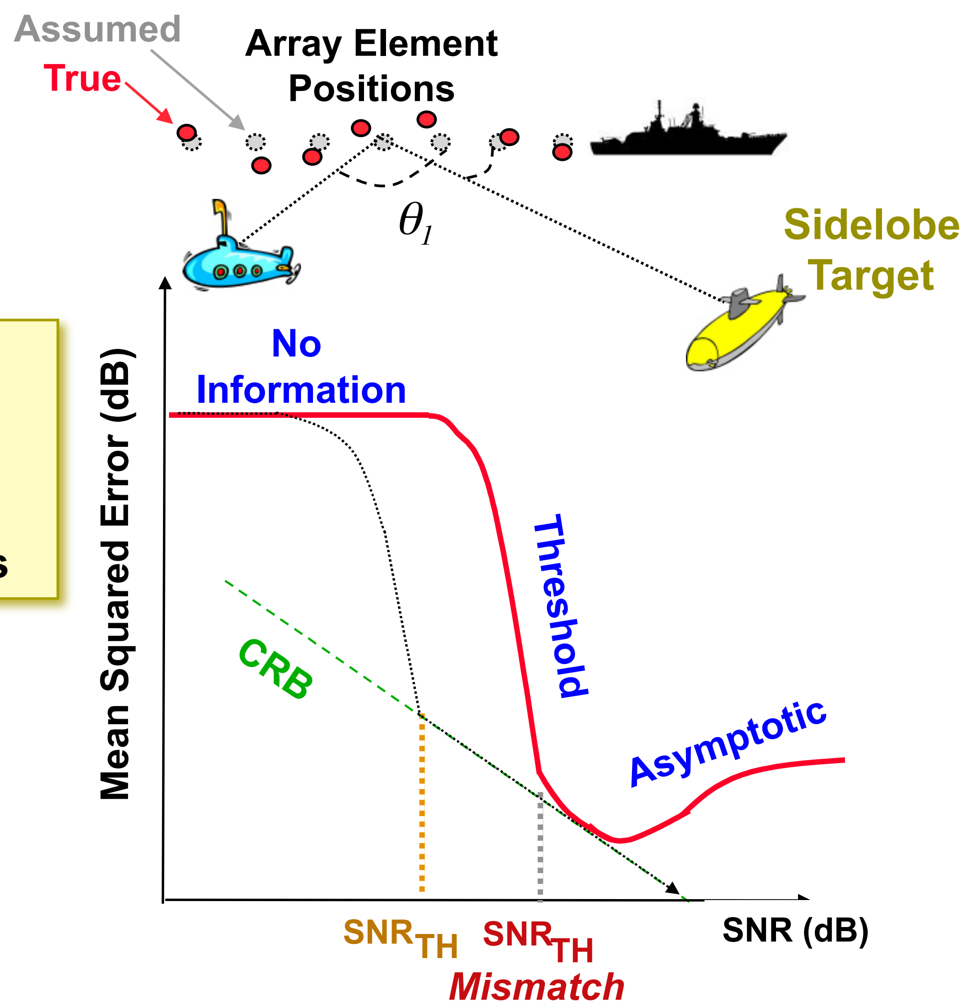
# Mean-Squared Error Performance: No Mismatch vs Mismatch

## No Mismatch




*Mismatch* affects threshold and asymptotic region leading to atypical performance curves

## Signal Mismatch



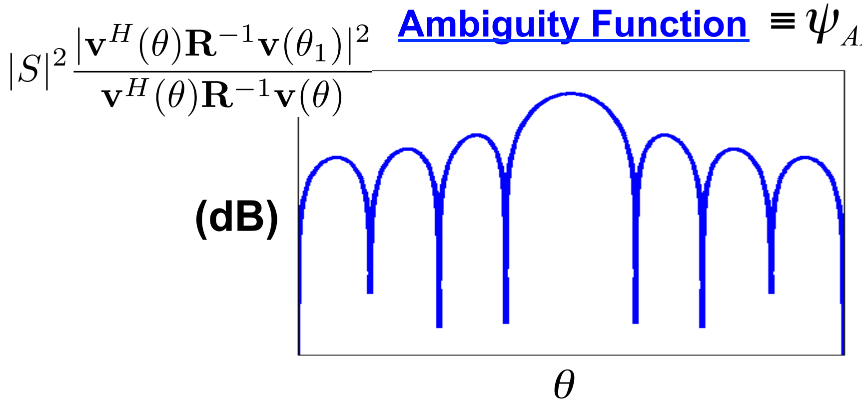
# Outline

- Introduction
- **Mean-squared error prediction**
- Numerical examples
- Summary

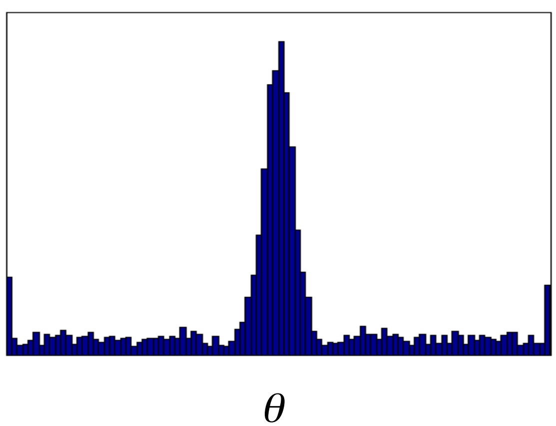
# Nonlinear Estimation and Ambiguity Functions

$$\widehat{\theta}_{ML} = \arg\max_{\theta} \frac{|\mathbf{v}^H(\theta)\widehat{\mathbf{R}}^{-1}\mathbf{x}|^2}{\mathbf{v}^H(\theta)\widehat{\mathbf{R}}^{-1}\mathbf{v}(\theta)}$$

**ML Ambiguity Function** $\equiv \psi_{AF}(\theta)$




**Histogram of ML Estimates**



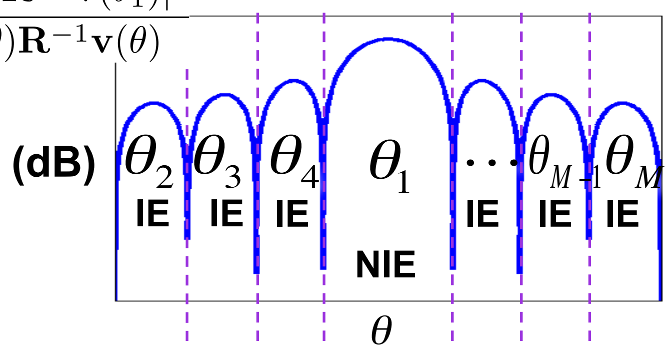

- **ML involves nonlinear search on finite interval**
- **PDF aggregates density around maxima of ambiguity function (AF)**
- **Divide interval into M sub-intervals**
  - **"No Interval Error" (NIE)**
    ***Local Errors***
  - **"Intervals of Error" (IE)**
    ***Global Errors***
- **Use intervals to approximate MSE**
  - **Estimation approximately a M-ary hypothesis testing problem**

# Nonlinear Estimation and Ambiguity Functions

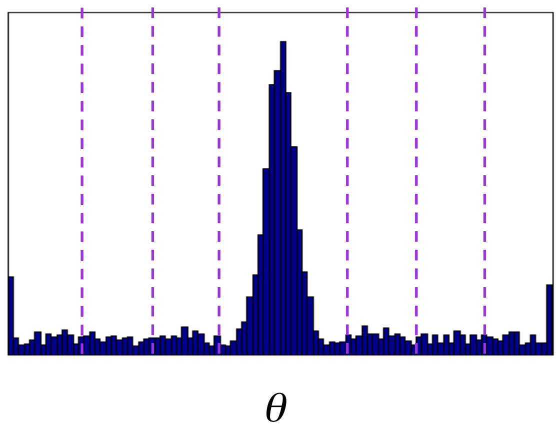
$$\widehat{\theta}_{ML} = \arg\max_{\theta} \frac{|\mathbf{v}^H(\theta)\widehat{\mathbf{R}}^{-1}\mathbf{x}|^2}{\mathbf{v}^H(\theta)\widehat{\mathbf{R}}^{-1}\mathbf{v}(\theta)}$$

**ML Ambiguity Function** $\equiv \psi_{AF}(\theta)$

$$|S|^2 \frac{|\mathbf{v}^H(\theta)\mathbf{R}^{-1}\mathbf{v}(\theta_1)|^2}{\mathbf{v}^H(\theta)\mathbf{R}^{-1}\mathbf{v}(\theta)}$$




**Histogram of ML Estimates**



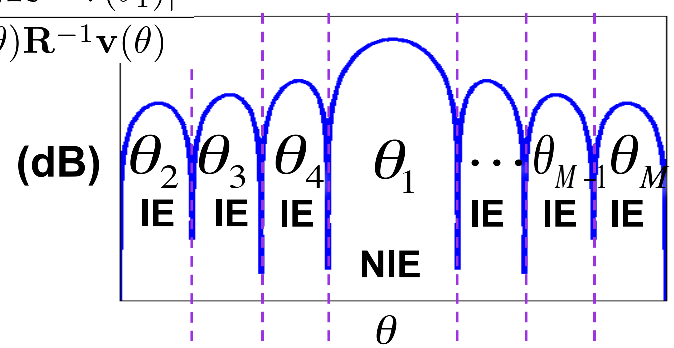

- **ML involves nonlinear search on finite interval**
- **PDF aggregates density around maxima of ambiguity function (AF)**
- **Divide interval into M sub-intervals**
  - **“No Interval Error” (NIE)**
    - ***Local Errors***
  - **“Intervals of Error” (IE)**
    - ***Global Errors***
- **Use intervals to approximate MSE**
  - **Estimation approximately a M-ary hypothesis testing problem**

# Nonlinear Estimation and Ambiguity Functions

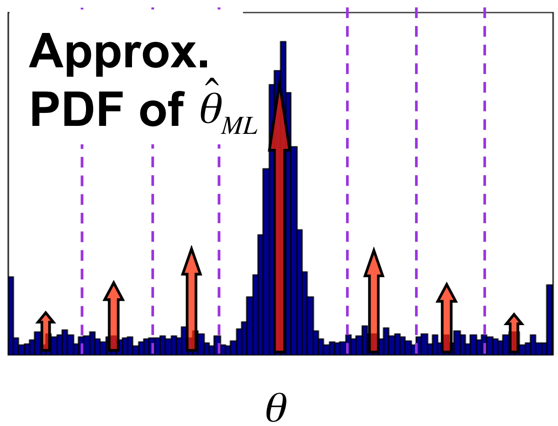
$$\widehat{\theta}_{ML} = \arg\max_{\theta} \frac{|\mathbf{v}^H(\theta)\widehat{\mathbf{R}}^{-1}\mathbf{x}|^2}{\mathbf{v}^H(\theta)\widehat{\mathbf{R}}^{-1}\mathbf{v}(\theta)}$$

**ML Ambiguity Function** $\equiv \psi_{AF}(\theta)$

$$|S|^2 \frac{|\mathbf{v}^H(\theta)\mathbf{R}^{-1}\mathbf{v}(\theta_1)|^2}{\mathbf{v}^H(\theta)\mathbf{R}^{-1}\mathbf{v}(\theta)}$$




**Histogram of ML Estimates**




- **ML involves nonlinear search on finite interval**
- **PDF aggregates density around maxima of ambiguity function (AF)**
- **Divide interval into M sub-intervals**
  - **"No Interval Error" (NIE)**
    - ***Local Errors***
  - **"Intervals of Error" (IE)**
    - ***Global Errors***
- **Use intervals to approximate MSE**
  - **Estimation approximately a M-ary hypothesis testing problem**

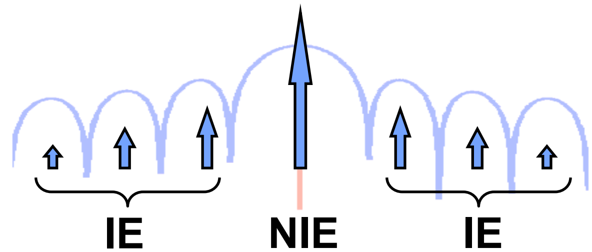
# Approximating MSE Performance: Based on Method of Interval Errors (MIE)

- **MSE given by** $E\left\{\left(\hat{\theta}-\theta_1\right)^2\right\} \equiv \int (\omega-\theta_1)^2\, p_{\hat{\theta}}(\omega)\, d\omega$

# Approximating MSE Performance: Based on Method of Interval Errors (MIE)

- **MSE given by** $$E\left\{\left(\hat{\theta}-\theta_1\right)^2\right\} \equiv \int (\omega-\theta_1)^2\, p_{\hat{\theta}}(\omega)\, d\omega$$

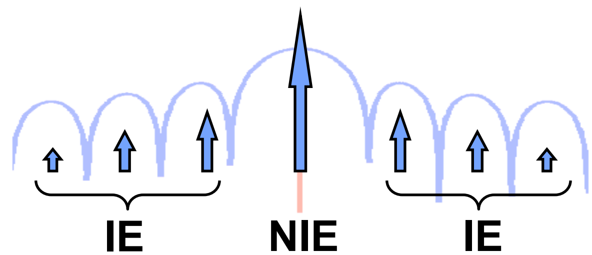
$$\approx \int (\omega-\theta_1)^2 \times$$




$$d\omega$$

# Approximating MSE Performance: Based on Method of Interval Errors (MIE)

- **MSE given by** $$E\left\{\left(\hat{\theta}-\theta_1\right)^2\right\} \equiv \int \left(\omega-\theta_1\right)^2 p_{\hat{\theta}}(\omega)\,d\omega$$

$$\approx \int \left(\omega-\theta_1\right)^2 \times$$  $$d\omega$$



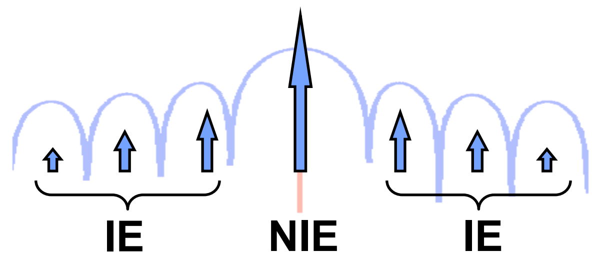
***"Local Errors"*** ***"Global Errors"***

$$\boxed{E\left\{\left(\hat{\theta}_{ML}-\theta_1\right)^2\middle|\theta_1\right\} \cong \left[1-\sum_{m=2}^{M} p\left(\hat{\theta}_{ML}=\theta_m\middle|\theta_1\right)\right]\cdot \sigma_{ML}^2\left(\theta_1\right) + \sum_{m=2}^{M} p\left(\hat{\theta}_{ML}=\theta_m\middle|\theta_1\right)\cdot\left(\theta_m-\theta_1\right)^2}$$

# Approximating MSE Performance: Based on Method of Interval Errors (MIE)

- **MSE given by** $$E\left\{\left(\hat{\theta}-\theta_1\right)^2\right\} \equiv \int \left(\omega-\theta_1\right)^2 p_{\hat{\theta}}(\omega)\,d\omega$$

$$\approx \int \left(\omega-\theta_1\right)^2 \times$$




$$d\omega$$

*"Local Errors"* *"Global Errors"*

$$E\left\{\left(\hat{\theta}_{ML}-\theta_1\right)^2 \middle| \theta_1\right\} \cong \left[1-\sum_{m=2}^{M} p\left(\hat{\theta}_{ML}=\theta_m \middle| \theta_1\right)\right] \cdot \sigma_{ML}^2\left(\theta_1\right) + \sum_{m=2}^{M} p\left(\hat{\theta}_{ML}=\theta_m \middle| \theta_1\right) \cdot \left(\theta_m-\theta_1\right)^2$$

- **Challenge is calculation of error probabilities and asymptotic MSE:**

$$p\left(\hat{\theta}_{ML}=\theta_m \middle| \theta_1\right) = ? \qquad \sigma_{ML}^2\left(\theta_1\right) = ?$$

***Both are functions of the estimated covariance***

# Interval Error Probability for Adaptive ML Estimation

- **Union bound suggests pairwise error approximation*:**

$$p\left(\widehat{\theta}_{ML} = \theta_m | E\{\mathbf{x}\} = S\mathbf{d}\right) \simeq$$

$$\Pr\left( \frac{|\mathbf{v}^H(\theta_m)\widehat{\mathbf{R}}^{-1}\mathbf{x}|^2}{\mathbf{v}^H(\theta_m)\widehat{\mathbf{R}}^{-1}\mathbf{v}(\theta_m)} > \frac{|\mathbf{v}^H(\theta_1)\widehat{\mathbf{R}}^{-1}\mathbf{x}|^2}{\mathbf{v}^H(\theta_1)\widehat{\mathbf{R}}^{-1}\mathbf{v}(\theta_1)} \middle| E\{\mathbf{x}\} = S\mathbf{d} \right)$$

**No Mismatch**

$$\mathbf{v}(\theta_1) = \mathbf{d}$$

**Array Response Mismatch**

$$\mathbf{v}(\theta_1) \neq \mathbf{d}$$

- **Asymptotic MSE approximated:**
  - **Adjustment of Cramér-Rao Bound (CRB)*** ***[No Mismatch]***
  - **Taylor series** ***[Array Response Mismatch]***



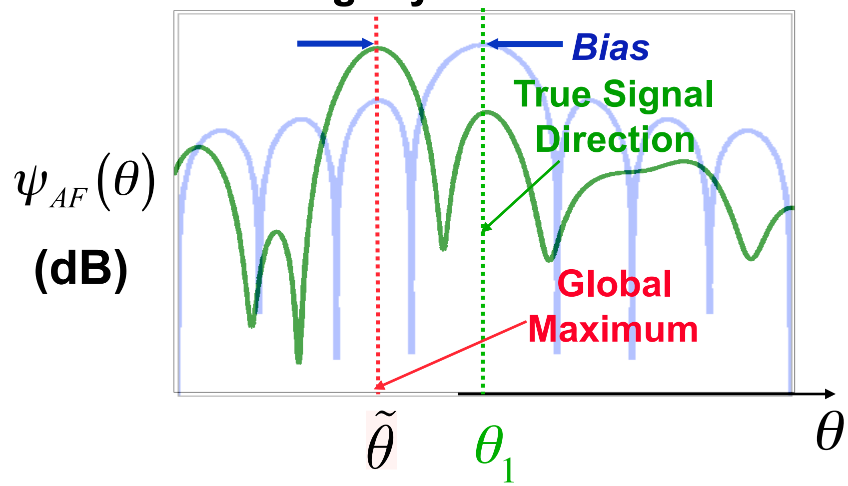
****Exact formulae derived in Richmond, IEEE T-IT, Vol. 52, No. 5, May 2006***

# Asymptotic ML MSE with Signal Mismatch

- **Asymptotic MSE (jitter) given by**

$$\sigma_{ML}^2(\theta_1) = E\left\{\left(\hat{\theta}_{ML} - \theta_1\right)^2\right\}$$

**ML Ambiguity Fnc with Mismatch**



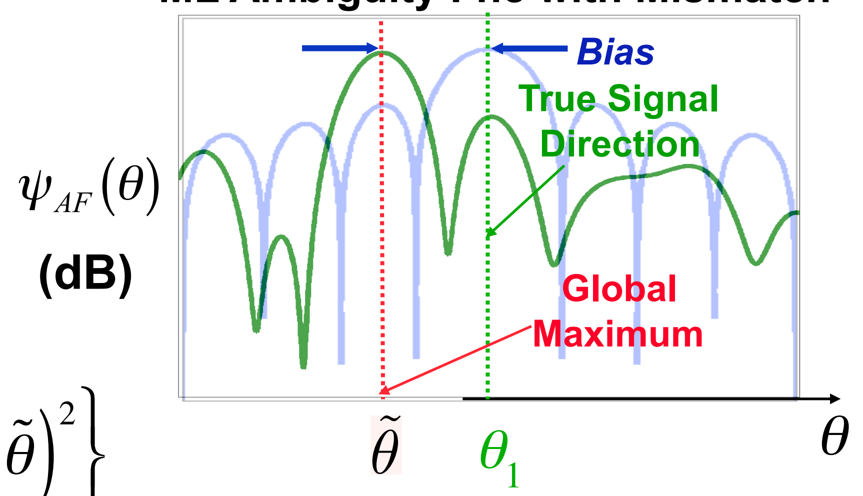
# Asymptotic ML MSE with Signal Mismatch

- **Asymptotic MSE (jitter) given by**

$$\sigma_{ML}^2(\theta_1) = E\left\{\left(\hat{\theta}_{ML} - \theta_1\right)^2\right\}$$

$$= \underbrace{\left[E\left\{\left(\hat{\theta}_{ML} - \theta_1\right)\right\}\right]^2}_{\textit{Bias}} + \underbrace{E\left\{\left(\hat{\theta}_{ML} - \tilde{\theta}\right)^2\right\}}_{\textit{Variance}}$$

**ML Ambiguity Fnc with Mismatch**




**where bias can be written as** $E\left\{\left(\widehat{\theta}_{ML} - \theta_1\right)\right\} = E\left\{\left(\widehat{\theta}_{ML} - \tilde{\theta}\right)\right\} + (\tilde{\theta} - \theta_1)$

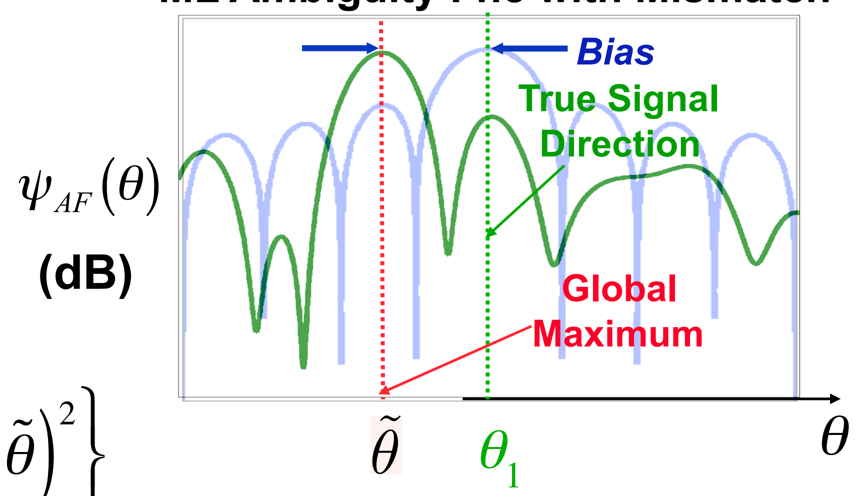
# Asymptotic ML MSE with Signal Mismatch

- **Asymptotic MSE (jitter) given by**

$$\sigma^2_{ML}(\theta_1) = E\left\{\left(\hat{\theta}_{ML} - \theta_1\right)^2\right\}$$

$$= \underbrace{\left[E\left\{\left(\hat{\theta}_{ML} - \theta_1\right)\right\}\right]^2}_{\textit{Bias}} + \underbrace{E\left\{\left(\hat{\theta}_{ML} - \tilde{\theta}\right)^2\right\}}_{\textit{Variance}}$$

**ML Ambiguity Fnc with Mismatch**




**where bias can be written as** $E\left\{\left(\hat{\theta}_{ML} - \theta_1\right)\right\} = E\left\{\left(\hat{\theta}_{ML} - \tilde{\theta}\right)\right\} + (\tilde{\theta} - \theta_1)$

Goal is to determine the necessary expectations

# ML Asymptotic MSE: Taylor Series Based Approach (1/3)

- **Define function** $f(\theta, \mathbf{A}, \mathbf{B}) = \dfrac{\mathbf{v}^H(\theta)\mathbf{A}\mathbf{B}\mathbf{A}^H\mathbf{v}(\theta)}{\mathbf{v}^H(\theta)\mathbf{A}\mathbf{v}(\theta)}$ **and note that**

$\widehat{\theta}_{ML} = \arg\max_{\theta} f(\theta, \widehat{\mathbf{R}}^{-1}, \widehat{\mathbf{R}}_T)$ **where** $\widehat{\mathbf{R}} \propto \sum_{l=1}^{L} \mathbf{x}(l)\mathbf{x}^H(l)$ **and** $\widehat{\mathbf{R}}_T = \mathbf{x}\mathbf{x}^H$

# ML Asymptotic MSE: Taylor Series Based Approach (1/3)

- **Define function** $f(\theta, \mathbf{A}, \mathbf{B}) = \dfrac{\mathbf{v}^H(\theta)\mathbf{A}\mathbf{B}\mathbf{A}^H\mathbf{v}(\theta)}{\mathbf{v}^H(\theta)\mathbf{A}\mathbf{v}(\theta)}$ **and note that**

$\widehat{\theta}_{ML} = \arg\max_{\theta} f(\theta, \widehat{\mathbf{R}}^{-1}, \widehat{\mathbf{R}}_T)$ **where** $\widehat{\mathbf{R}} \propto \sum_{l=1}^{L} \mathbf{x}(l)\mathbf{x}^H(l)$ **and** $\widehat{\mathbf{R}}_T = \mathbf{x}\mathbf{x}^H$

- **Define partials** $\dot{f}(\theta, \mathbf{A}, \mathbf{B}) = \dfrac{\partial f(\theta, \mathbf{A}, \mathbf{B})}{\partial \theta}$ **and** $\ddot{f}(\theta, \mathbf{A}, \mathbf{B}) = \dfrac{\partial^2 f(\theta, \mathbf{A}, \mathbf{B})}{\partial^2 \theta}$

# ML Asymptotic MSE: Taylor Series Based Approach (1/3)

- **Define function** $f(\theta, \mathbf{A}, \mathbf{B}) = \dfrac{\mathbf{v}^H(\theta)\mathbf{A}\mathbf{B}\mathbf{A}^H\mathbf{v}(\theta)}{\mathbf{v}^H(\theta)\mathbf{A}\mathbf{v}(\theta)}$ **and note that**

$\widehat{\theta}_{ML} = \arg\max_{\theta} f(\theta, \widehat{\mathbf{R}}^{-1}, \widehat{\mathbf{R}}_T)$ **where** $\widehat{\mathbf{R}} \propto \sum_{l=1}^{L} \mathbf{x}(l)\mathbf{x}^H(l)$ **and** $\widehat{\mathbf{R}}_T = \mathbf{x}\mathbf{x}^H$

- **Define partials** $\dot{f}(\theta, \mathbf{A}, \mathbf{B}) = \dfrac{\partial f(\theta, \mathbf{A}, \mathbf{B})}{\partial \theta}$ **and** $\ddot{f}(\theta, \mathbf{A}, \mathbf{B}) = \dfrac{\partial^2 f(\theta, \mathbf{A}, \mathbf{B})}{\partial^2 \theta}$

- **Taylor's theorem allows the first order approximation:**

$$\dot{f}(\widehat{\theta}_{ML}, \widehat{\mathbf{R}}^{-1}, \widehat{\mathbf{R}}_T) \simeq \dot{f}(\tilde{\theta}, \mathbf{R}^{-1}, \mathbf{R}_T) + \ddot{f}(\tilde{\theta}, \mathbf{R}^{-1}, \mathbf{R}_T)(\widehat{\theta}_{ML} - \tilde{\theta})$$

$$+\mathrm{tr}\left[\left(\frac{\partial \dot{f}}{\partial \mathbf{B}}\right)^T_{\mathbf{B}=\mathbf{R}_T}\left(\widehat{\mathbf{R}}_T - \mathbf{R}_T\right)\right] + \mathrm{tr}\left[\left(\frac{\partial \dot{f}}{\partial \mathbf{A}}\right)^T_{\mathbf{A}=\mathbf{R}^{-1}}\left(\widehat{\mathbf{R}}^{-1} - \mathbf{R}^{-1}\right)\right] + \mathrm{h.o.t.}$$

**where** $\tilde{\theta}$ **is the argument obtaining the peak value of** $f(\theta, \mathbf{R}^{-1}, \mathbf{R}_T)$

# ML Asymptotic MSE:
# Taylor Series Based Approach (1/3)

- **Define function** $f(\theta, \mathbf{A}, \mathbf{B}) = \dfrac{\mathbf{v}^H(\theta)\mathbf{A}\mathbf{B}\mathbf{A}^H\mathbf{v}(\theta)}{\mathbf{v}^H(\theta)\mathbf{A}\mathbf{v}(\theta)}$ **and note that**

$\widehat{\theta}_{ML} = \arg\max_{\theta} f(\theta, \widehat{\mathbf{R}}^{-1}, \widehat{\mathbf{R}}_T)$ **where** $\widehat{\mathbf{R}} \propto \sum_{l=1}^{L} \mathbf{x}(l)\mathbf{x}^H(l)$ **and** $\widehat{\mathbf{R}}_T = \mathbf{x}\mathbf{x}^H$

- **Define partials** $\dot{f}(\theta, \mathbf{A}, \mathbf{B}) = \dfrac{\partial f(\theta, \mathbf{A}, \mathbf{B})}{\partial\theta}$ **and** $\ddot{f}(\theta, \mathbf{A}, \mathbf{B}) = \dfrac{\partial^2 f(\theta, \mathbf{A}, \mathbf{B})}{\partial^2\theta}$

- **Taylor's theorem allows the first order approximation:**

$$\underbrace{\dot{f}(\widehat{\theta}_{ML}, \widehat{\mathbf{R}}^{-1}, \widehat{\mathbf{R}}_T)}_{\text{Vanish}} \simeq \underbrace{\dot{f}(\tilde{\theta}, \mathbf{R}^{-1}, \mathbf{R}_T)}_{\text{Vanish}} + \ddot{f}(\tilde{\theta}, \mathbf{R}^{-1}, \mathbf{R}_T)(\widehat{\theta}_{ML} - \tilde{\theta})$$

$$+\mathrm{tr}\left[\left(\frac{\partial \dot{f}}{\partial \mathbf{B}}\right)^T_{\mathbf{B}=\mathbf{R}_T}\left(\widehat{\mathbf{R}}_T - \mathbf{R}_T\right)\right] + \mathrm{tr}\left[\left(\frac{\partial \dot{f}}{\partial \mathbf{A}}\right)^T_{\mathbf{A}=\mathbf{R}^{-1}}\left(\widehat{\mathbf{R}}^{-1} - \mathbf{R}^{-1}\right)\right] + \mathrm{h.o.t.}$$

**where** $\tilde{\theta}$ **is the argument obtaining the peak value of** $f(\theta, \mathbf{R}^{-1}, \mathbf{R}_T)$

# ML Asymptotic MSE: Taylor Series Based Approach (2/3)

- **Since** $\dot{f}(\widehat{\theta}_{ML}, \widehat{\mathbf{R}}^{-1}, \widehat{\mathbf{R}}_T) = 0$ **and** $\dot{f}(\tilde{\theta}, \mathbf{R}^{-1}, \mathbf{R}_T) = 0$ **the MSE can be expressed as**

$$\widehat{\theta}_{ML} - \tilde{\theta} \simeq \frac{1}{\ddot{f}(\tilde{\theta}, \mathbf{R}^{-1}, \mathbf{R}_T)} \cdot \left( \mathrm{tr}\left[ \left( \frac{\partial \dot{f}}{\partial \mathbf{B}} \right)^T_{\mathbf{B}=\mathbf{R}_T} \Delta \mathbf{R}_T \right] + \mathrm{tr}\left[ \left( \frac{\partial \dot{f}}{\partial \mathbf{A}} \right)^T_{\mathbf{A}=\mathbf{R}^{-1}} \Delta \mathbf{R}^{-1} \right] \right)$$

**where** $\Delta \mathbf{R}_T = \widehat{\mathbf{R}}_T - \mathbf{R}_T$ **and** $\Delta \mathbf{R}^{-1} = \widehat{\mathbf{R}}^{-1} - \mathbf{R}^{-1}$

# ML Asymptotic MSE: Taylor Series Based Approach (2/3)

- **Since** $\dot{f}(\widehat{\theta}_{ML}, \widehat{\mathbf{R}}^{-1}, \widehat{\mathbf{R}}_T) = 0$ **and** $\dot{f}(\tilde{\theta}, \mathbf{R}^{-1}, \mathbf{R}_T) = 0$ **the MSE can be expressed as**

$$\widehat{\theta}_{ML} - \tilde{\theta} \simeq \frac{1}{\ddot{f}(\tilde{\theta}, \mathbf{R}^{-1}, \mathbf{R}_T)} \cdot \left( \mathrm{tr}\left[ \left( \frac{\partial \dot{f}}{\partial \mathbf{B}} \right)^T_{\mathbf{B}=\mathbf{R}_T} \Delta\mathbf{R}_T \right] + \mathrm{tr}\left[ \left( \frac{\partial \dot{f}}{\partial \mathbf{A}} \right)^T_{\mathbf{A}=\mathbf{R}^{-1}} \Delta\mathbf{R}^{-1} \right] \right)$$

**where** $\Delta\mathbf{R}_T = \widehat{\mathbf{R}}_T - \mathbf{R}_T$ **and** $\Delta\mathbf{R}^{-1} = \widehat{\mathbf{R}}^{-1} - \mathbf{R}^{-1}$

- **Note that** $E\left\{ \left( \widehat{\theta}_{ML} - \tilde{\theta} \right) \right\} = 0$ **since** $E\left\{ \Delta\mathbf{R}_T \right\} = \mathbf{0}$ **and normalization of** $\widehat{\mathbf{R}}^{-1}$ **can be chosen such that** $E\left\{ \Delta\mathbf{R}^{-1} \right\} = \mathbf{0}$

# ML Asymptotic MSE: Taylor Series Based Approach (2/3)

- Since $\dot{f}(\widehat{\theta}_{ML}, \widehat{\mathbf{R}}^{-1}, \widehat{\mathbf{R}}_T) = 0$ and $\dot{f}(\tilde{\theta}, \mathbf{R}^{-1}, \mathbf{R}_T) = 0$ the MSE can be expressed as

$$\widehat{\theta}_{ML} - \tilde{\theta} \simeq \frac{1}{\ddot{f}(\tilde{\theta}, \mathbf{R}^{-1}, \mathbf{R}_T)} \cdot \left( \mathrm{tr}\left[ \left( \frac{\partial \dot{f}}{\partial \mathbf{B}} \right)^T_{\mathbf{B}=\mathbf{R}_T} \Delta \mathbf{R}_T \right] + \mathrm{tr}\left[ \left( \frac{\partial \dot{f}}{\partial \mathbf{A}} \right)^T_{\mathbf{A}=\mathbf{R}^{-1}} \Delta \mathbf{R}^{-1} \right] \right)$$

where $\Delta \mathbf{R}_T = \widehat{\mathbf{R}}_T - \mathbf{R}_T$ and $\Delta \mathbf{R}^{-1} = \widehat{\mathbf{R}}^{-1} - \mathbf{R}^{-1}$

- Note that $E\left\{ \left( \widehat{\theta}_{ML} - \tilde{\theta} \right) \right\} = 0$ since $E\left\{ \Delta \mathbf{R}_T \right\} = \mathbf{0}$ and normalization of $\widehat{\mathbf{R}}^{-1}$ can be chosen such that $E\left\{ \Delta \mathbf{R}^{-1} \right\} = \mathbf{0}$

- Mismatch analysis addressed via choice of

$$\mathbf{R}_T = E\{\mathbf{x}\mathbf{x}^H\} = \mathbf{R} + |S|^2 \mathbf{d}\mathbf{d}^H \quad \text{such that} \quad \mathbf{v}(\theta_1) \neq \mathbf{d}$$

# ML Asymptotic MSE: Taylor Series Based Approach (3/3)

- **Squaring quadratic, and taking expectation one obtains the desired asymptotic MSE:**

$$E\left\{\left(\widehat{\theta}_{ML}-\tilde{\theta}\right)^2\right\} \simeq \frac{E\ \mathrm{tr}^2\left[\left(\dfrac{\partial \dot{f}}{\partial \mathbf{B}}\right)^T_{\mathbf{B}=\mathbf{R}_T} \Delta\mathbf{R}_T\right]}{[\ddot{f}(\tilde{\theta},\mathbf{R}^{-1},\mathbf{R}_T)]^2} + \frac{E\ \mathrm{tr}^2\left[\left(\dfrac{\partial \dot{f}}{\partial \mathbf{A}}\right)^T_{\mathbf{A}=\mathbf{R}^{-1}} \Delta\mathbf{R}^{-1}\right]}{[\ddot{f}(\tilde{\theta},\mathbf{R}^{-1},\mathbf{R}_T)]^2}$$



**Richmond, Proceedings of IEEE SAM Workshop, 2006*

# ML Asymptotic MSE: Taylor Series Based Approach (3/3)

- **Squaring quadratic, and taking expectation one obtains the desired asymptotic MSE:**

$$E\left\{\left(\widehat{\theta}_{ML}-\tilde{\theta}\right)^2\right\} \simeq \frac{E\,\mathrm{tr}^2\left[\left(\frac{\partial \dot{f}}{\partial \mathbf{B}}\right)^T_{\mathbf{B}=\mathbf{R}_T} \Delta\mathbf{R}_T\right]}{[\ddot{f}(\tilde{\theta},\mathbf{R}^{-1},\mathbf{R}_T)]^2} + \frac{E\,\mathrm{tr}^2\left[\left(\frac{\partial \dot{f}}{\partial \mathbf{A}}\right)^T_{\mathbf{A}=\mathbf{R}^{-1}} \Delta\mathbf{R}^{-1}\right]}{[\ddot{f}(\tilde{\theta},\mathbf{R}^{-1},\mathbf{R}_T)]^2}$$

***NEW***

$\simeq$ **MSE R known*** **+** **Accuracy Loss due to Unknown R**

- **MSE is that obtained assuming $\mathrm{R}^{-1}$ is known plus an additional term accounting for accuracy loss due to estimation of $\mathrm{R}^{-1}$**



***Richmond, Proceedings of IEEE SAM Workshop, 2006**

# ML Asymptotic MSE: Taylor Series Based Approach (3/3)

- **Squaring quadratic, and taking expectation one obtains the desired asymptotic MSE:**

$$E\left\{\left(\widehat{\theta}_{ML}-\tilde{\theta}\right)^2\right\} \simeq \frac{E\,\mathrm{tr}^2\left[\left(\frac{\partial \dot{f}}{\partial \mathbf{B}}\right)^T_{\mathbf{B}=\mathbf{R}_T}\Delta\mathbf{R}_T\right]}{[\ddot{f}(\tilde{\theta},\mathbf{R}^{-1},\mathbf{R}_T)]^2} + \frac{E\,\mathrm{tr}^2\left[\left(\frac{\partial \dot{f}}{\partial \mathbf{A}}\right)^T_{\mathbf{A}=\mathbf{R}^{-1}}\Delta\mathbf{R}^{-1}\right]}{[\ddot{f}(\tilde{\theta},\mathbf{R}^{-1},\mathbf{R}_T)]^2}$$

***NEW***

$\simeq$ **MSE R known*** **+** **Accuracy Loss due to Unknown R**

- **MSE is that obtained assuming $\mathrm{R}^{-1}$ is known plus an additional term accounting for accuracy loss due to estimation of $\mathrm{R}^{-1}$**
- **Moments of complex Gaussian process, and moments of inverted Wishart complete analysis**

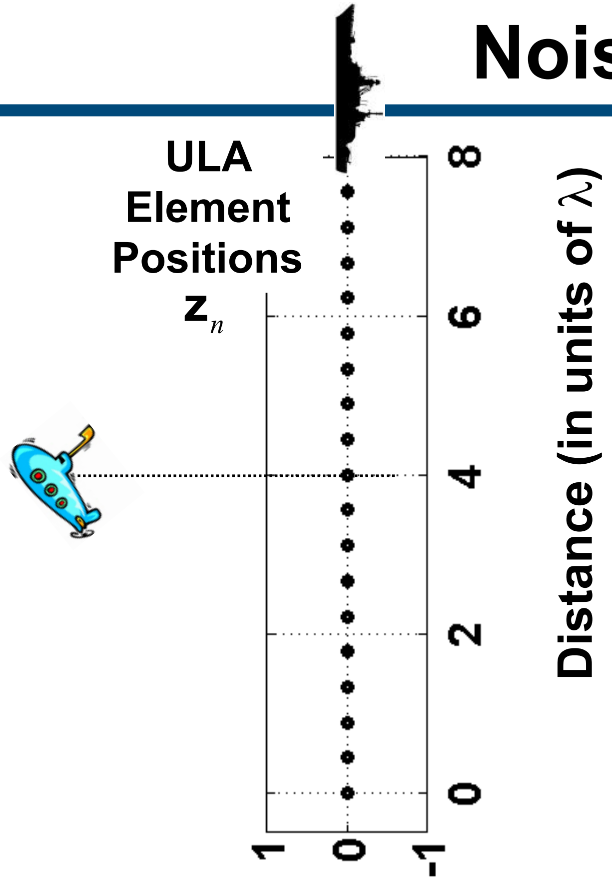


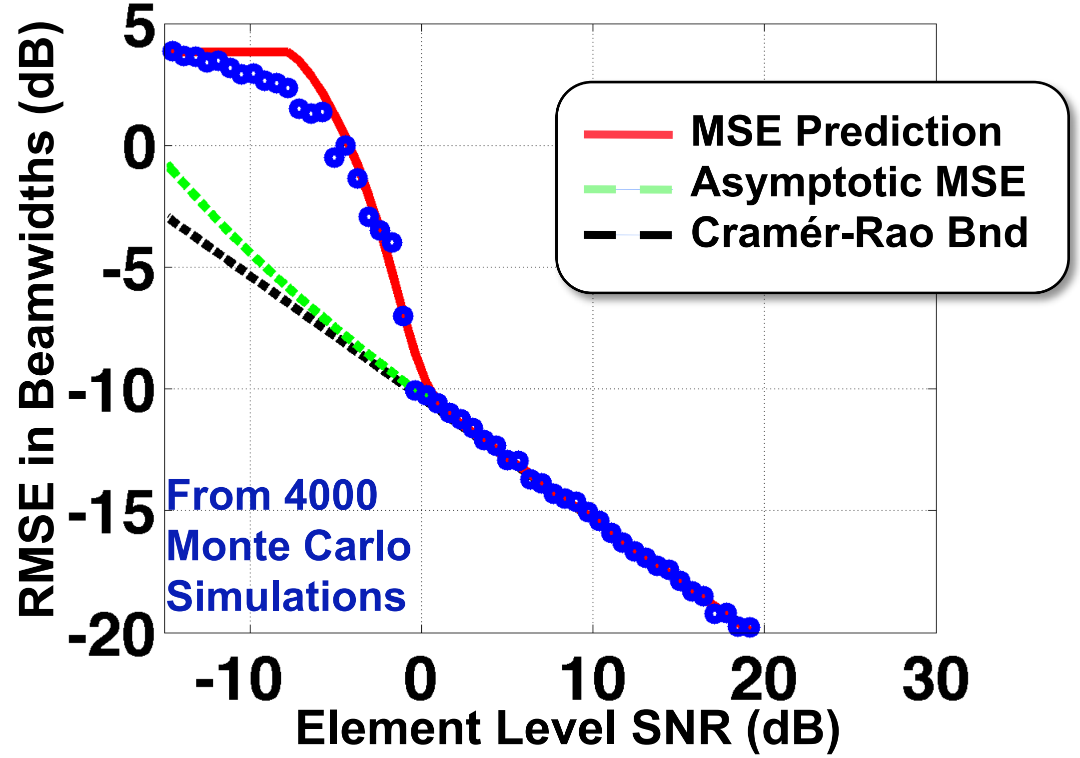
**Richmond, Proceedings of IEEE SAM Workshop, 2006*

# Outline

- Introduction
- Mean-squared error prediction
- ✓ **Numerical examples**
- Summary

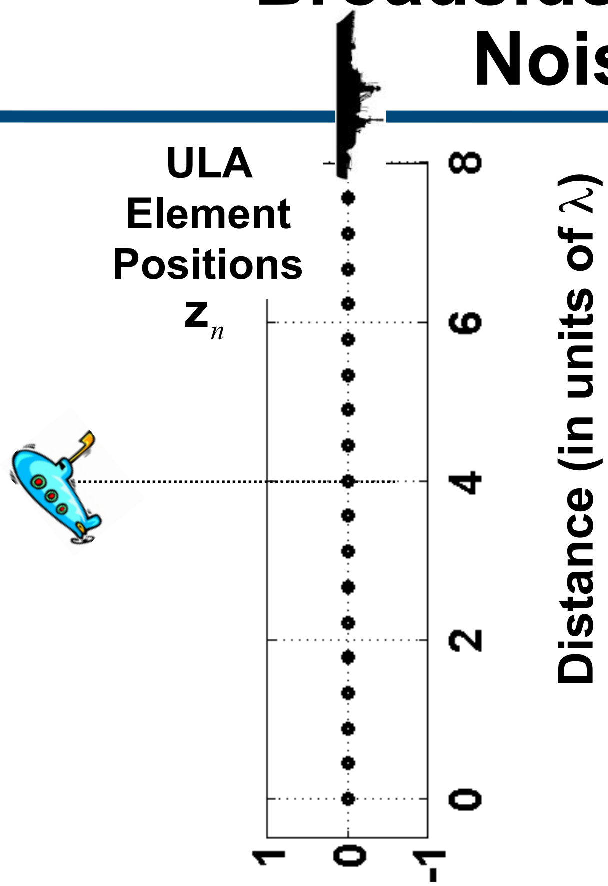
# Broadside Planewave Signal in White Noise: No Mismatch, ULA

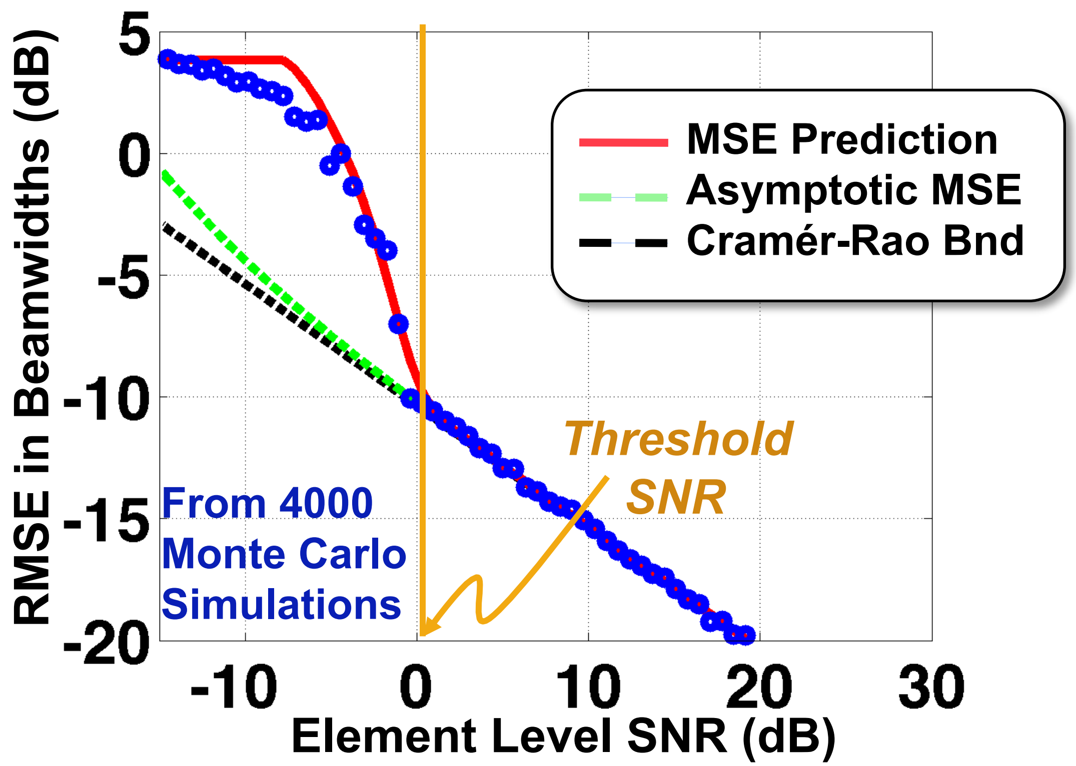



- $N=18$ **element uniform linear array (ULA),** $(\lambda/2.25)$ **element spacing**
  - **3dB Beamwidth ≈ 7.2 degs, search space [60 120] degs**
  - **0dB white noise, True Signal @ 90 degs (broadside)**
- **Asymptotic ML MSE agrees with CRB above threshold SNR**
- **MIE MSE predictions very accurate above and below threshold**

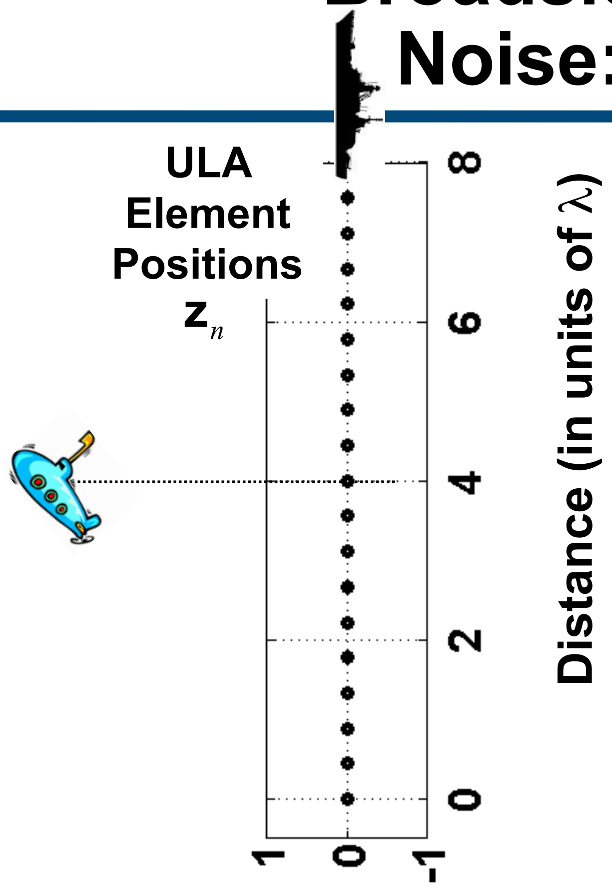
# Broadside Planewave Signal in White Noise: No Mismatch, ULA

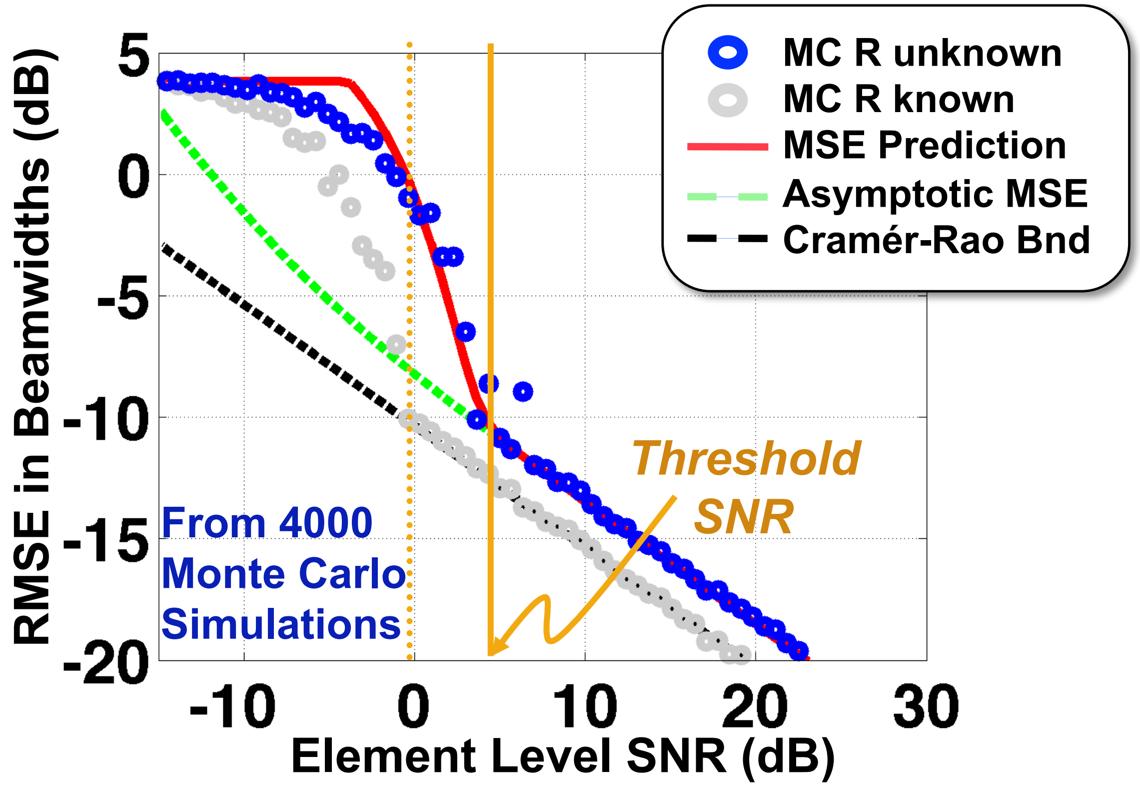



- $N$=18 element uniform linear array (ULA), ($\lambda$/2.25) element spacing
  - 3dB Beamwidth ≈ 7.2 degs, search space [60 120] degs
  - 0dB white noise, True Signal @ 90 degs (broadside)
- Asymptotic ML MSE agrees with CRB above threshold SNR
- MIE MSE predictions very accurate above and below threshold

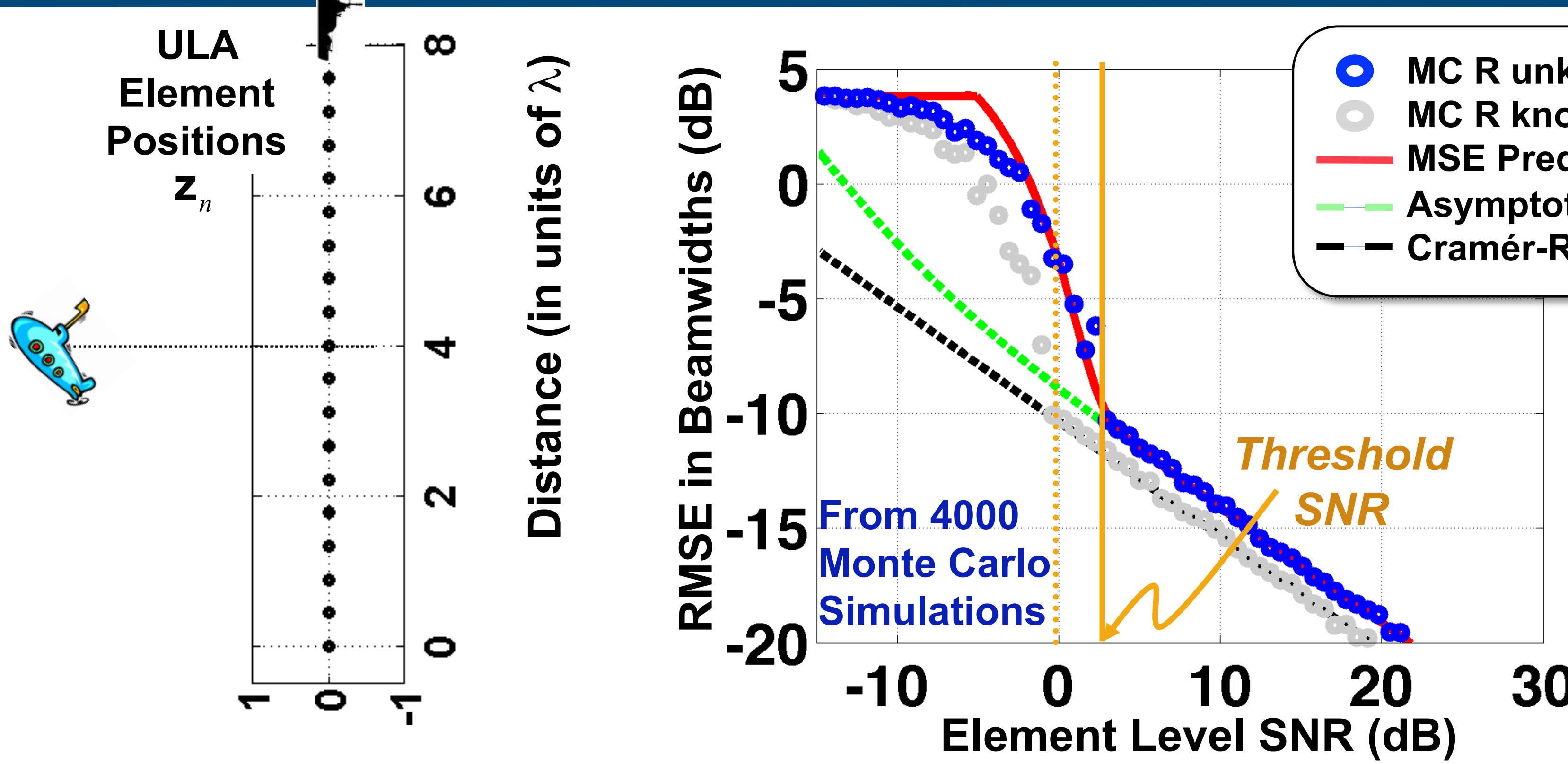
# Broadside Planewave Signal in White Noise: No Mismatch, ULA, *L=1.5N*



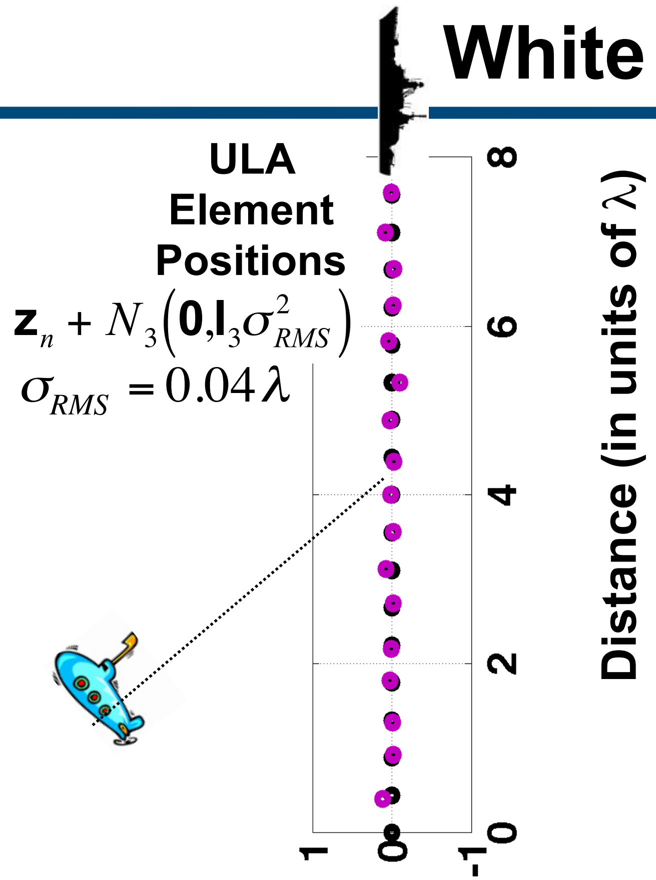

- $N$=18 element uniform linear array (ULA), $(\lambda/2.25)$ element spacing
  - 3dB Beamwidth ≈ 7.2 degs, search space [60 120] degs
  - 0dB white noise, True Signal @ 90 degs (broadside)
- Asymptotic ML MSE agrees with CRB above threshold SNR
- MIE MSE predictions very accurate above and below threshold

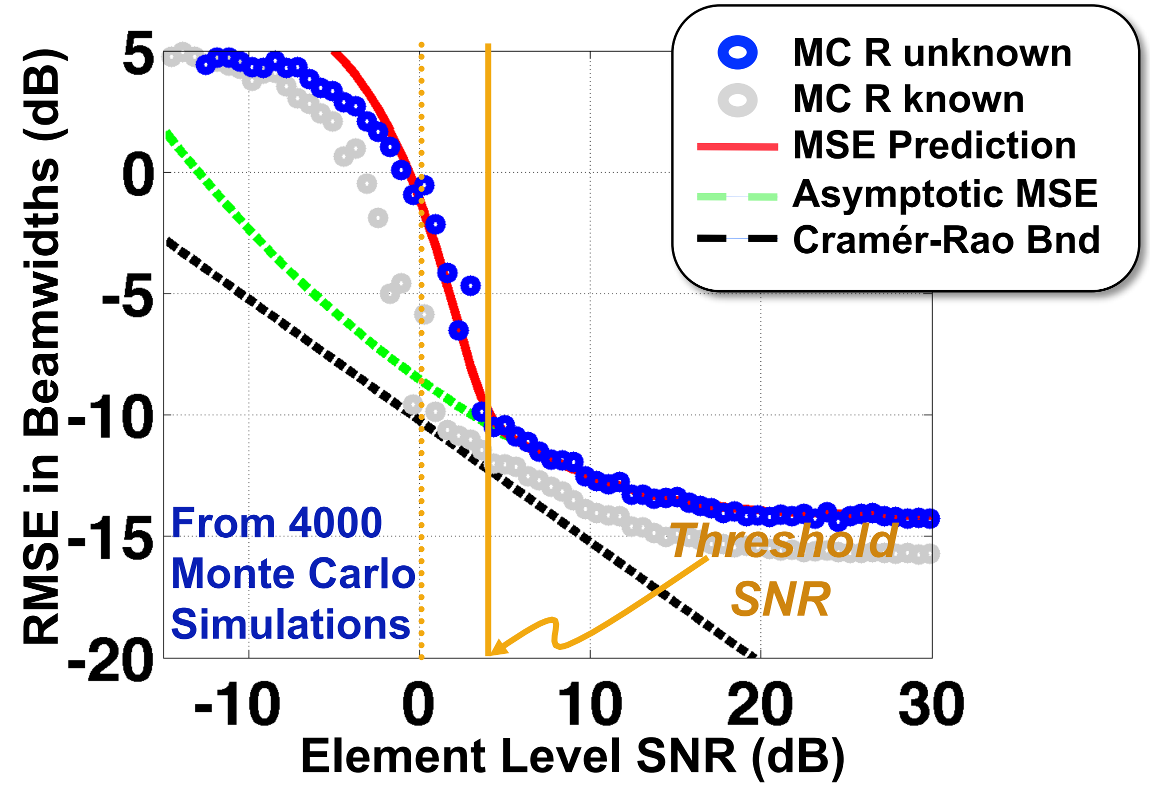
# Broadside Planewave Signal in White Noise: No Mismatch, ULA, *L=2N*




- $N=18$ **element uniform linear array (ULA),** $(\lambda/2.25)$ **element spacing**
  - **3dB Beamwidth ≈ 7.2 degs, search space [60 120] degs**
  - **0dB white noise, True Signal @ 90 degs (broadside)**
- **Asymptotic ML MSE agrees with CRB above threshold SNR**
- **MIE MSE predictions very accurate above and below threshold**

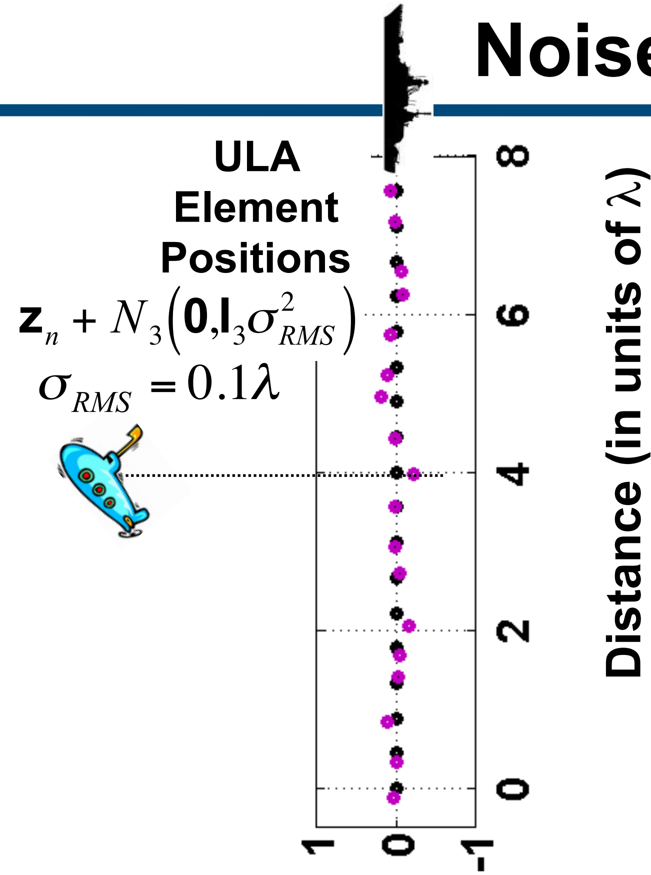
# Off Broadside Planewave Signal in White Noise: Perturbed ULA, *L=2N*

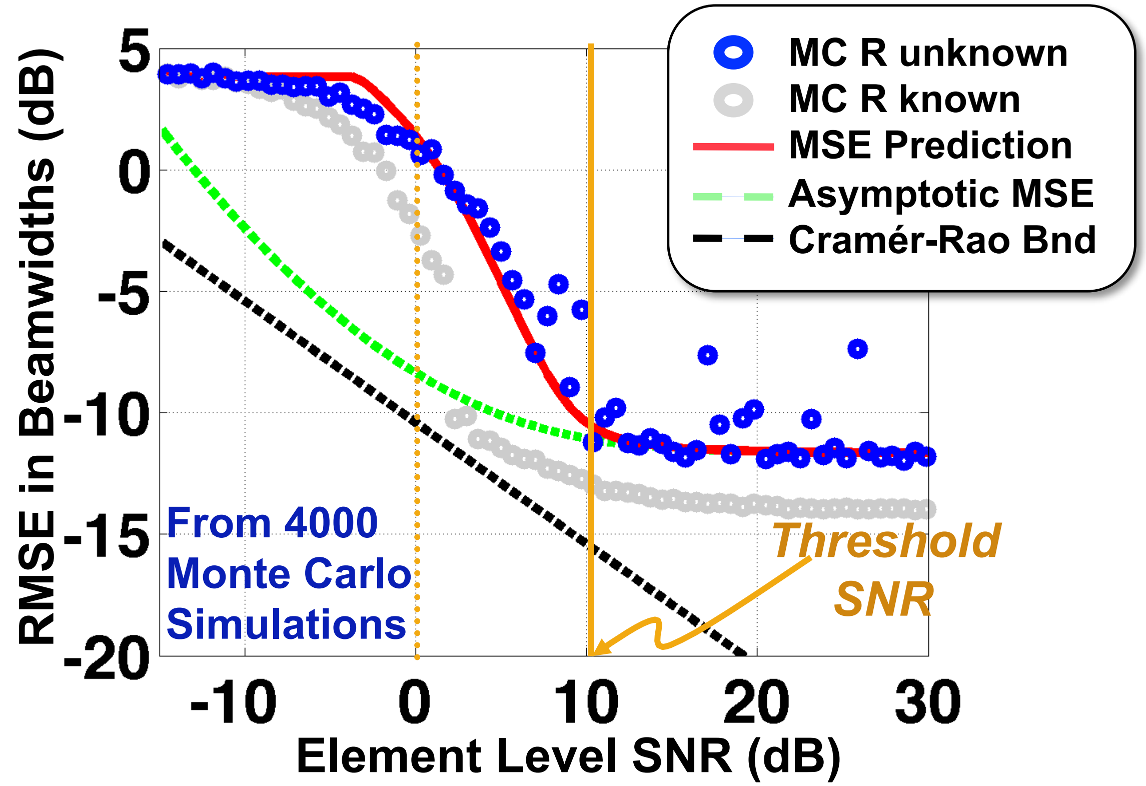



- $N{=}18$ **element ULA positions perturbed by 3-D Gaussian noise**
  - ***Zero mean with stand. dev.*** $0.04\lambda$
  - ***Array perturbation from single realization of Gaussian noise process***
- **CRB no longer useful, while MIE accurate above and below threshold**
- **Mismatch can increase threshold SNR and introduce asymptotic bias**

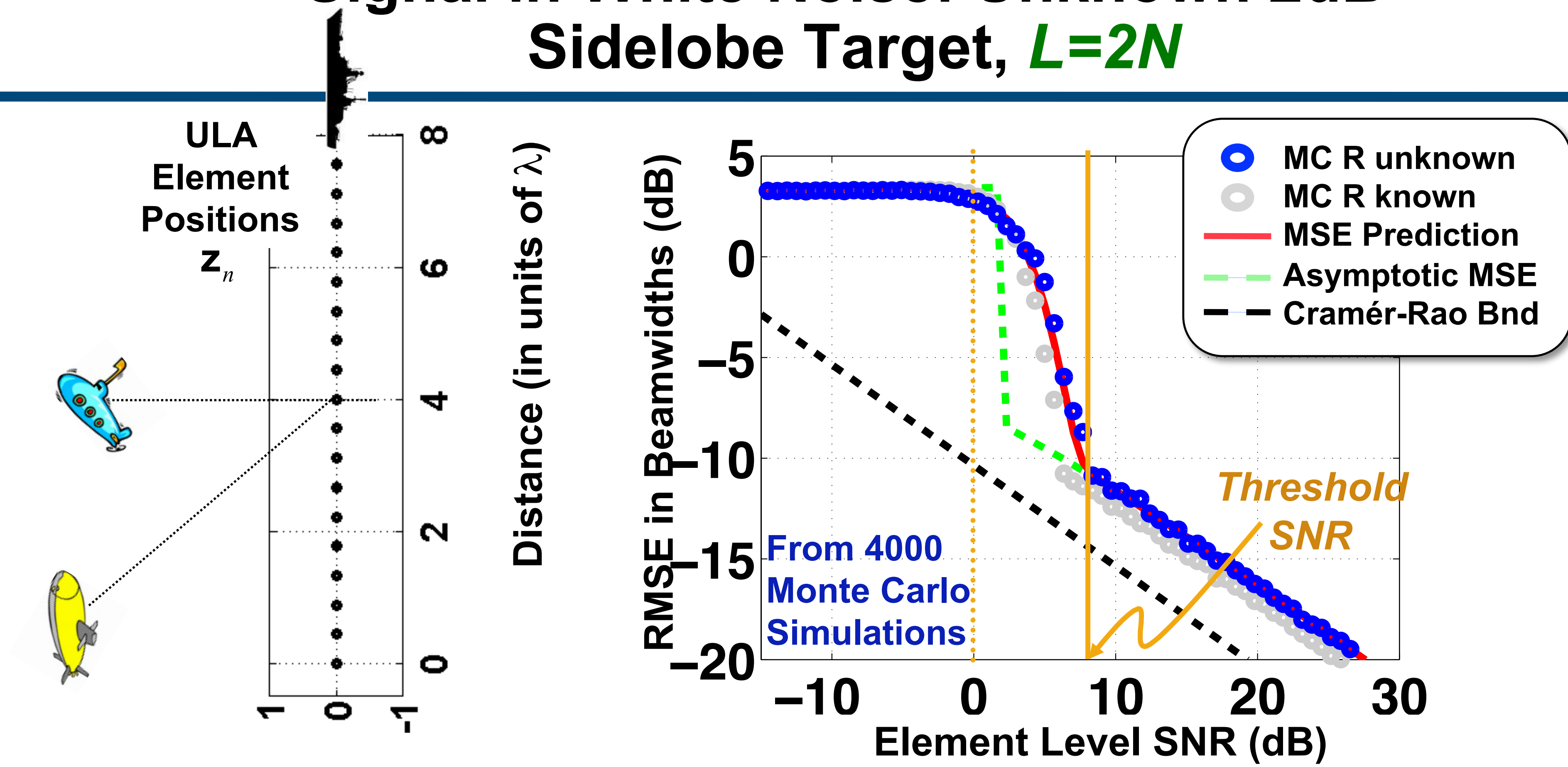
# Broadside Planewave Signal in White Noise: Perturbed ULA, *L=2N*



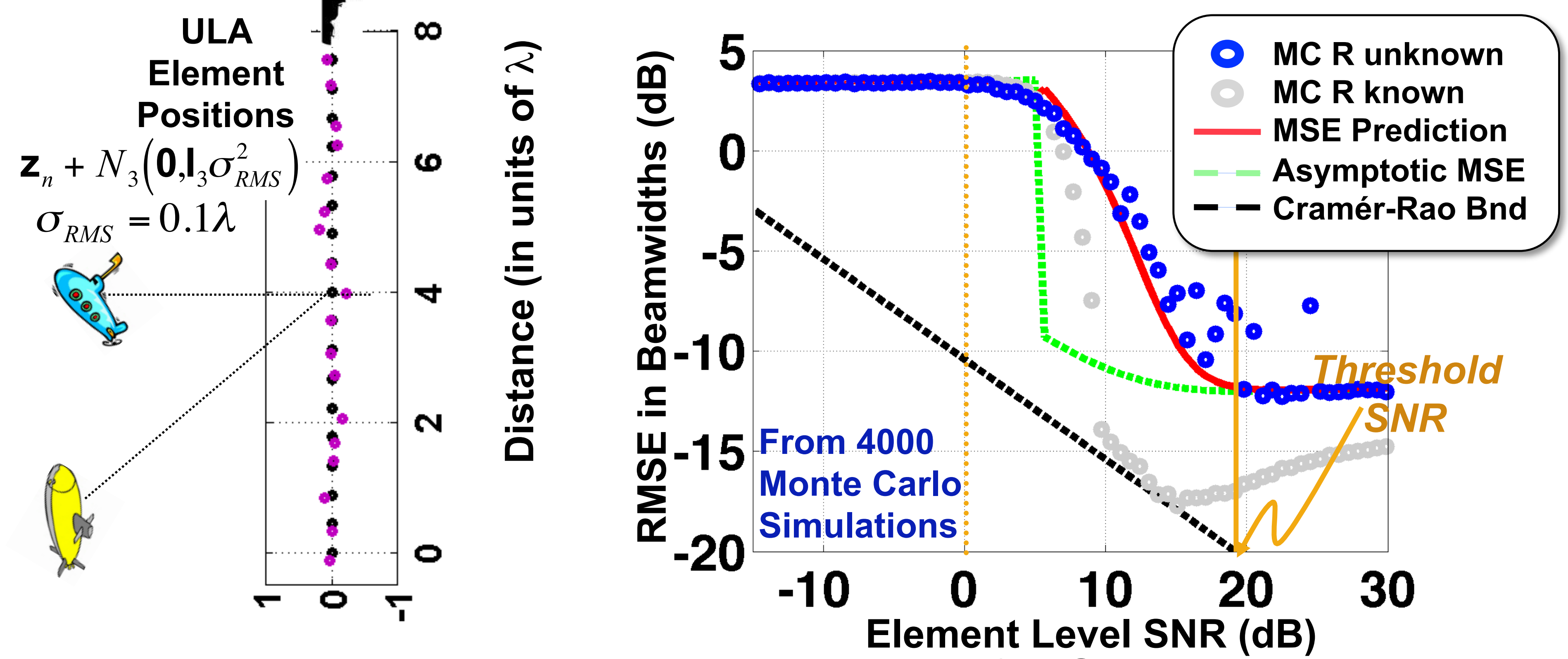

- $N$=18 element ULA positions perturbed by 3-D Gaussian noise
  - *Zero mean with stand. dev.* $0.1\lambda$
  - *Array perturbation from single realization of Gaussian noise process*
- CRB no longer useful, while MIE accurate above and below threshold
- Mismatch can increase threshold SNR and introduce asymptotic bias

# Signal in White Noise: Unknown 2dB Sidelobe Target, *L=2N*




- $N=18$ **element uniform linear array (ULA),** $(\lambda/2.25)$ **element spacing**
  - *Sidelobe target 2dB above noise @ 75degs*
- **SINR loss and threshold region results from competing targets**
- **Asymptotic MSE ∝ 1/SNR due to signal living on ML manifold**

# Signal in White Noise: Perturbed ULA + 2dB Sidelobe Target (SLT), *L=2N*




- $N=18$ **element ULA positions perturbed by 3-D Gaussian noise**
  - *Zero mean with stand. dev.* $0.1\lambda$
  - *Array perturbation from single realization of Gaussian noise process*
  - *Sidelobe target 2dB above noise @ 75degs*
- **CRB no longer useful, while MIE accurate above and below threshold**
- **Mismatch affects both threshold region and asymptotic performance**

# Outline

- Introduction
- Mean-squared error prediction
- Numerical examples
- **Summary**

# Summary

- Developed theory to predict impact of signal mismatch on ML estimation accuracy
  - arbitrary deterministic signal mismatch allowed
  - above and below threshold SNR performance considered
  - finite sample effects due to noise covariance estimation included
- Generalizes previous work on ML threshold region performance and provides tools useful for system design and analysis
- Circumvents need for time consuming Monte Carlo simulation
- Examples show estimation loss introduced by uncertainties in R and v can be significant and should be considered in design

# Backups

# Abstract

## Asymptotic Mean Squared Error Performance of Maximum-Likelihood DOA Estimation with Estimated Noise Covariance

Christ D. Richmond*
*Senior Member, IEEE*, christ@ll.mit.edu

The mean squared error (MSE) performance prediction of Maximum-Likelihood (ML) Direction-Of-Arrival (DOA) angle estimation has been studied extensively by several authors (see refs in [1]). Most recently ML DOA performance prediction of both the threshold and asymptotic regions in the presence of a general form of deterministic array response mismatch was considered assuming the noise-plus-interference covariance matrix (NICM) was known [1,2]. The error probabilities required for predicting the threshold region MSE of the general case of deterministic mismatch are derived in [3] including the unknown NICM case. Approximations of the asymptotic MSE performance were explored in [2] based on a stochastic representation of the ML filter weight vector, yielding only moderate success due to the approximate representation of the functional dependence of some parameters. Herein an exact general expression for the asymptotic DOA MSE performance for the unknown NICM case is derived based on a Taylor Series expansion that accounts for the exact functional dependence of all parameters. It is shown herein that the MSE expression derived in [1] is augmented by an additive term that accounts for the loss due exclusively to NICM estimation.

[1] C. D. Richmond, "On the Threshold Region Mean-Squared Error Performance of Maximum-Likelihood Direction-Of-Arrival Estimation in the Presence of Signal Model Mismatch," *Proceedings of the Fourth IEEE SAM Processing Workshop*, pp. 268—272, Waltham, MA, July 2006.

[2] C. D. Richmond, "Signal Model Mismatch and Maximum-Likelihood Mean-Squared Error Performance," *Proceedings of the Adaptive Sensor Array Processing Workshop*, MIT Lincoln Laboratory, June 2006.

[3] C. D. Richmond, "Mean Squared Error and Threshold SNR Prediction of Maximum-Likelihood Signal Parameter Estimation with Estimated Colored Noise Covariances," *IEEE Transactions on Information Theory*, Vol. 52, No. 5, pp. 2146—2164, May 2006.

# ML Asymptotic MSE: Approximate Taylor Series Based Approach*

- Following Taylor Series approach it can be shown that

$$\hat{\theta}_{ML} - \theta_G \approx -\frac{2/K}{\ddot{f}(\theta_G, \mathbf{R}_T, \mathbf{0}, 1/K)}\Big\{\underbrace{\mathrm{Re}\big[\dot{\mathbf{w}}^H(\theta_G)\Delta\mathbf{R}_T\mathbf{w}(\theta_G)\big]}_{\text{Known R}} +$$

$$\underbrace{\mathrm{Re}\big[\dot{\mathbf{w}}^H(\theta_G)\mathbf{R}_T\mathbf{t}\big] + \mathrm{Re}\big[\dot{\mathbf{w}}^H(\theta_G)\mathbf{R}_T\mathbf{w}(\theta_G)\big](\chi - 1/K)\Big\}}_{\text{Unknown R}}$$

MSE obtained from Mean and Variance of $\hat{\theta}_{ML} - \theta_G$

- Quadratic function defined $f(\theta, \hat{\mathbf{R}}_T, \mathbf{t}, \chi) \equiv \chi[\mathbf{w}(\theta) + \mathbf{t}]^H \hat{\mathbf{R}}_T [\mathbf{w}(\theta) + \mathbf{t}]$

where $\mathbf{w}(\theta) \equiv \dfrac{\mathbf{R}^{-1}\mathbf{v}(\theta)}{\sqrt{\mathbf{v}^H(\theta)\mathbf{R}^{-1}\mathbf{v}(\theta)}}$, $\hat{\mathbf{w}}(\theta) \equiv \dfrac{\hat{\mathbf{R}}^{-1}\mathbf{v}(\theta)}{\sqrt{\mathbf{v}^H(\theta)\hat{\mathbf{R}}^{-1}\mathbf{v}(\theta)}} = \chi[\mathbf{w}(\theta) + \mathbf{t}]$

$$\hat{\mathbf{R}}_T = \mathbf{x}\mathbf{x}^H, \qquad E\{\mathbf{x}\mathbf{x}^H\} \equiv \mathbf{R}_T = \mathbf{R} + |S|^2 \mathbf{d}\mathbf{d}^H, \qquad \hat{\mathbf{R}} \equiv \frac{1}{L}\mathbf{X}\mathbf{X}^H$$

$\chi$ and t are random quantities* modeling errors in w($\theta$) due to unknown R



*Richmond, Proceedings of ASAP 2006*